The Biometric Society

BIOMETRICS

FOUNDED BY THE BIOMETRICS SECTION OF THE AMERICAN STATISTICAL ASSOCIATION

TABLE OF CONTENTS

Number 2 **June 1957** **Volume 13**

The Biometric Society is an international society devoted to the mathematical and statistical aspects of biology. Biologists, mathematicians, statisticians and others interested in its objectives are invited to become members. Through its regional organizations the Society sponsors regional and local meetings. National secretaries serve the interests of members in Denmark, India, Japan, the Netherlands, Sweden and Switzerland, and there are many members at large.

Rates (in U.S.A. currency) for full membership in the Society for 1957, including dues and a subscription to this journal are: for residents of Canada and the United States $7.00; for others $4.50. Members of the American Statistical Association who are currently subscribing to *Biometrics* through that organization may become members of the Biometric Society on payment of $3.00 annual dues if resident in the United States or Canada, and of $1.75 annual dues if resident elsewhere. Information concerning the Society can be obtained from its Secretary, M. J. R. Healy, Statistics Department, Rothamsted Experimental Station, Harpenden, Herts., England.

Annual subscription for non-members of the Biometric Society is $7.00, payable to the Managing Editor, *Biometrics*, National Research Council, Ottawa 2, Canada. Members of the American Statistical Association may subscribe through the Secretary of the American Statistical Association at $4.00 per annum.

Second-class mailing privileges authorized at Knoxville, Tenn. Additional entry at Richmond, Va. Business Office: 509 West Hill Road, Knoxville 19, Tennessee. *Biometrics* is published quarterly—in March, June, September and December.

RANDOMIZED BRANCH SAMPLING FOR ESTIMATING FRUIT NUMBER

S. C. Pearce and D. A. Holland
East Malling Research Station, Kent, England

A method called "randomized branch sampling" has recently been advanced (Jessen [1955]) for determining the number of fruits on a tree. In this paper its usefulness will be considered, especially in comparison with a long-established method (Hoblyn [1931]) available for complete enumeration and readily adapted to sampling.

The particular form of randomized branch sampling recommended was that denoted by the symbol *PPA*, *i.e.*, "probabilities proportional to area". The recorder starts at the trunk and measures the girths of the main branches; he calculates their squares and selects a branch randomly, the probability of selection being proportional to basal area. He follows up the branch chosen to the next fork, where he repeats the process, and so on. After a number of such selections he arrives at a unit of the tree upon which he counts the fruit, and an estimate of fruit number for the whole tree is then calculated.

Such a method of sampling can be judged only in relation to the purpose of the investigation and the species concerned. Thus, the purpose may be the estimation of the crop on only the particular tree or set of trees being sampled, as in an experiment, or the trees may themselves constitute a sample, as when a survey is made of crops in a large area; different standards of accuracy might be required in the various cases.

Again, with some species the branches fork more frequently than with others, and there may be differences in ease of access to the forks for measurement, and there may be varying chances of causing damage while doing so. Also, some species usually bear their fruit on new wood, some on old, and some on both. Consequently it is unlikely that any one method can be used in all instances without modification.

The claim is made that the procedure recommended is operationally simple. No doubt it would be possible to minimize the amount of arithmetic by using a nomogram or table to give $a^2/(a^2 + b^2)$ from a and b, and a selection of random numbers could be written down

beforehand. Nevertheless, the experience of the present writers with a range of European fruit tree species leads them to question this operational simplicity, at least as a general rule. Also, the result of sampling is not readily checked unless marks are made on the tree to show the result of each random selection.

However, even if practicable, the method does not appear to be very accurate. Data are given from only one tree, a pineapple orange in Florida in 1953, with results that are rather discouraging. This tree had two main branches, one of which had three sub-branches and the other had two. When these five sub-branches were taken successively as units for counting, it was found that they severally gave estimates of the total number of fruit on the tree as follows: 1696, 903, 874, 1701, 1171, the true figure being 1379 and the variance of the error of estimation 128,545. For some purposes this would not be accurate enough. Thus, the coefficient of variability of crops of mature trees is commonly about 30–40 per cent and may be less (Hoblyn [1931]), so for purposes of an experiment with one-tree plots it is desirable that the standard error introduced by sampling shall not exceed 10 per cent of the mean if its effect on the overall accuracy is to be negligible, *i.e.*, the variance needs to be less than about 19,000.

Some interest attaches to the assumption that the number of fruit on a branch is proportional to the square of the branch girth. The question does not appear to have been considered previously; though Wilcox [1940–41], working with apples, has suggested that the weight of fruit on a tree varies as $(\text{trunk girth})^{1.5}$. On general grounds, however, it is to be expected that the quantity of fruit would be directly related to the weight of the branch, and unpublished work at East Malling on apples suggests that this varies as the branch girth raised to the power 2.8 or 3.0, which supports the findings of Sudds and Anthony [1929]. Certainly, for the tree from which data are available, substitution of cubes for squares leads to the estimates: 1625, 1117, 1164, 1427, 1194, which are better, the variance being now 37,931. However, the use of the fourth power is yet better, the estimates becoming 1587, 1373, 1635, 1235, 1243, and the variance 26,994, but even these results are hardly good enough for the purpose mentioned, while the method appears arbitrary.

Although the method will be unbiassed whatever power is used, it is desirable to adopt a value that is justified biologically and this will on the average lead to the greatest accuracy. In the absence of a physiological basis the choice of power must be made empirically, but to do this justifiably would need an extensive study. A figure should not be accepted on the basis of a few data unless there were convincing

prior reasons for it. Judging from experience with the analogous problem of relating tree weight to girth (Pearce [1952]), it is not certain that any one power will be generally valid.

From the data presented there is little hope that precision would be improved by taking selection beyond the sub-branches and counting more units keeping the total sample about one-fifth of the tree; while the work of measuring girths, raising to a power and randomizing would be increased. Improvement would therefore have to come from increasing the sample size.

In this connection, it is worth recalling the method proposed some quarter-century ago for measuring the one-year-old shoots on trees (Hoblyn [1931]). This makes use of the fact that, at any fork, one branch must morphologically be more distal than the other, and consists simply of recording the more distal one first. This method has been found to work very satisfactorily for a number of species over many years not only for measuring shoots, incidence of disease, etc, but for counting fruits, blossoms, and leaves. For these last purposes a "clicker-counter" is invaluable. This method can be so rapid that in some instances no sampling method appears to be called for.

However, if one is needed, Hoblyn's method of enumeration, which in effect arranges shoots in an objective order, readily provides one, for it is often sufficient to observe every nth shoot instead of all of them. Such a method could reasonably be expected to be better than "*PPA* with sub-branches," because, according to Jessen's table, this is about as efficient as "*PE* with smallest branches", and the proposed method should be better than this for a given sample size, partly because it uses smaller sampling units and partly because it disperses the sample over the tree instead of taking it at random. A difficulty in testing this suggestion from Jessen's data arises from his giving no indication which branch at a fork is more distal, this information not being required by his method. Accordingly, ten examples have been considered, all based on the orange tree for which Jessen published his data. In five examples, the branches at each fork were given a random order, the first being accounted the most distal; in the others, the branches were given an order at random only where none was suggested by subsequent branchings as being botanically likely. For the completely random examples variances for $n = 3$, 5, and 10 were respectively 17,770, 15,005, and 14,502; for the partly random ones the corresponding figures were 113,120, 13,355, and 4,938. On account of the lack of data these figures should not be regarded as very reliable; but they are quite encouraging. It may be noted that Jessen's "smallest branches" are larger units than one-year shoots, and the latter should therefore give

more accurate estimates. On the other hand, some unit larger than one-year shoots might be preferable if fruit were partly borne on old wood. With Jessen's method, about one-fifth of the fruits must be counted and certain branch girths measured; with the method suggested here, for n equal to five the amount of counting will be about the same, no branch girths will be needed, but the whole tree will have to be surveyed, so the time spent will be about the same or perhaps less. The evidence suggests that accuracy will be better. Nevertheless, it is not clear what is the best method of recording; this is regrettable in view of the importance of the subject, and further work is needed. Until more is known the writers would prefer to keep to Hoblyn's method of complete enumeration, or, if that were too laborious, they would count the fruit on every n-th shoot.

SUMMARY

The method of randomised branch sampling for fruit number is examined and compared with some alternatives.

REFERENCES

Hoblyn, T. N. [1931]. Field experiments in horticulture. *Tech. Commun. Bur. Fruit Prod.* East Malling, *2*.

Jessen, Raymond, J. [1955]. Determining the fruit count on a tree by randomized branch sampling. *Biometrics, 11,* 99–109.

Pearce, S. C. [1952]. Studies in the measurement of apple trees. I. The use of trunk girth to estimate tree size. *Rep. E. Malling Res. Sta. for 1951,* 101–4.

Sudds, R. H. and Anthony, R. D. [1929]. The correlation of trunk measurements with tree performance in apples. *Proc. Amer. Soc. Hort. Sci., 25,* 244–6.

Wilcox, J. C. [1940–41]. Adjusting apple yields for differences in size of tree. *Sci. Agric., 21,* 139–48.

THE ERROR OF REPLICATED POTENCY ESTIMATES IN A BIOLOGICAL ASSAY METHOD OF THE PARALLEL LINE TYPE*

MINDEL C. SHEPS AND PAUL L. MUNSON

Department of Biostatistics, School of Public Health, Biological Research Laboratories and Departments of Preventive Medicine and Pharmacology, School of Dental Medicine and Medical School, Harvard University, Boston, Mass., U.S.A.

INTRODUCTION

In the practice of biological assay the most widely used formulas [2, 3] for the variance and the fiducial limits of a potency estimate take account of intra-assay error only. The assumption implicit in the use of these formulas that inter-assay variation need not be considered may be tested by comparing the variation predicted by the formula with the variation actually observed in replicated assays of unknowns. Such a test in a collaborative penicillin assay of the parallel line type analyzed by Bliss [4] showed that the actual error of the log potency estimate (M) was considerably and significantly greater than the error predicted from intra-assay statistics. It was suggested that a large portion of the discrepancy might be due to errors in dosage of test substances. A similar suggestion was made by Dews and Berkson [5] in their consideration of the actual error of quantal assays. Excess variability in M occurring during collaborative trials for International Standards of several antibiotics [6–10] was at times associated with, although not fully explained by, significant deviations from linearity or from parallelism. Although many of the observed discrepancies were small in an absolute sense, they were detected because of the considerable innate (within-group) precision of the assay methods. The actual error of M for several other International Standards [11–13] was also greater than predicted by the usual computations. Interactions of M with replicate samples of the same milk within assays and also between assays were observed by Clarke [14] in parallel line assays for riboflavin potency.

*Presented in part before the Eighth and Ninth (1) Annual Meetings of the Biometric Society, Eastern North American Region, New York, December 27, 1955 and Detroit, September 10, 1956.

This investigation was supported in part by research grants (C-2056 and C-2191) from the National Cancer Institute of the National Institutes of Health, Public Health Service.

These results and others like them [15–17] from both quantal and graded-response assay methods suggest that the occurrence of heterogeneity in a group of replicate M's may be due not only to occasional accidental gross errors in procedure, or to basic non-validity of the assay [18, 19], but also to random experimental variation. In such cases, the sources of inter-assay error must be sought, and measures taken to offset them. If inter-assay variation can be eradicated or progressively diminished by improvements in the method, the internal estimate of error is approached as a minimum. However, as long as inter-assay error is present, the precision of the method cannot be estimated adequately by the internal statistics alone and an accurate estimate of the total error of M including the inter-assay variation is needed.

We have investigated this problem in several rather large series of routinely performed parallel line assays in which many unknowns were tested repeatedly. The most extensive data were obtained from a chick comb weight assay method for androgens as performed over a period of about three years. A smaller series of replicate M's from a parathyroid assay method have also been analyzed.

SOURCE AND SELECTION OF ANGROGEN ASSAY DATA

Assay Procedure

The initial procedure for the androgen assays was based on the method of Rakoff et al. [20] with several modifications [21, 22]. Our experiments on the assay method have led to five successive modifications in procedure resulting in five series of similar assays. Series I and II were conducted with ether as the vehicle for androgen. They differed from each other in the volume of ether solution applied per dose and in the dosage of androsterone, the reference standard (Table 1). Series III was a group of experimental assays that will not be considered here. Ethanol was substituted for the more volatile ether as the vehicle for androgen in Series IV and V, which were distinguished from each other only by the dosage of androsterone administered.

Each assay included two or three dose levels of androsterone, and several unknowns at one or two dose levels. The unknowns were either extracts of human urine collected in the course of several clinical endocrine studies [21, 23–25], or synthetic steroids chemically different from androsterone. Each unknown was assayed at least twice. More than two tests were conducted when an entire assay was invalid, when there was notable disagreement between the first two potency estimates, or for increased precision. The time interval between the replicated

assays of a given unknown varied, and naturally tended to be longer for unknowns with more than two replications.

TABLE 1

CHARACTERISTICS OF THE ASSAY METHOD FOR ANDROGENS IN FOUR SERIES OF ASSAYS VARYING IN DESIGN AND TECHNIQUE

	Series I	Series II	Series IV	Series V
Number of assays	13	61	28	16
Vehicle for the androgen	Ether	Ether	Ethanol	Ethanol
Volume of solution used per application	0.05 ml.	0.01 ml.	0.01 ml.	0.01 ml.
Range in daily dose for the standard	0.1–1.6μg.	0.05–0.8μg.	0.05–0.8μg.	0.1–0.8μg.
Approximate number of chicks for entire assay	100	100	200	200
for the standard	30	30	40	40
Mean for the standard slope ($\bar{b}_s$)	24.2	24.2	24.2	30.8
Inter-assay variance of standard slope (observed) $V(b_s)$	36.6	30.0	15.4	18.4
Variance of standard slope estimated from intra-assay statistics	15	16	7	12
Mean residual variance per response	105	118	95	94
Mean $\lambda(\bar{s}/\bar{b})$	0.42	0.45	0.40	0.32
Median value of g in random sample of assays	0.04	0.04	0.01	0.01

Criteria for Selection of Data

All replicated assays that met certain definite criteria were included in this report. Data were used only from assays in which the responses to the reference standard were linear, with a slope lying within the 99% tolerance limits of the mean slope for standard in all assays of the series. Furthermore, within these acceptable assays, only those potency estimates were used that were based on responses parallel to the standard curve ($P > .01$). The distribution of 279 replicate M's that satisfied these criteria is shown in Table 2.*

Comparison of the Assay Series

The effect of the androgen on the chick comb, that is the assay response, was expressed as the coded logarithm of the ratio of the comb

*Tabulations of the replicate M's are not included in this report, but mimeographed copies of these data are available and will be supplied on request.

TABLE 2

DISTRIBUTION OF REPLICATE LOG POTENCY ESTIMATES FROM ANDROGEN ASSAYS

	Assay series I and II	Assay series IV and V		All
	Urine extracts	Urine extracts	Synthetic steroids	
Unknowns with:				
Two estimates	60	28	18	106
Three estimates	11	0	4	15
Four estimates	4	0	0	4
Six estimates	0	0	1	1
Total unknowns	75	28	23	126
Number of assays yielding replicates	62	16	18	96
Number of replicate estimates	169	56	54	279

weight to the body weight. The procedural and statistical characteristics of all the assays from Series I, II, IV and V are summarized separately for each series in Table 1. The intra-assay statistics from the first two series (ether assays) were similar. Although the index of precision ($\lambda = s/b$) was undesirably high, it was possible to compensate for it partially by the use, in each assay, of approximately 30 chicks for the standard and 10–20 chicks for each unknown.

In Series IV and V (ethanol assays) the number of chicks for androsterone was increased to approximately 40. The assays in these series were characterized by a slight reduction in residual error variance and an increased slope of the regression line, resulting in a moderate (25–30%) improvement in the index of precision, λ. Altogether, about 16,000 chicks were used in the assays reported here.

It would have been difficult to compare the combined slopes (b_c) from the individual assays because they included various combinations, at various dose levels, of some of the unknowns under consideration here with other unknowns. Instead, the standard slopes (b_s) were compared, since the dose levels and the number of chicks per dose level of the standard were essentially uniform throughout each series. Obviously, b_s for any assay was not greatly different from b_c based on parallel response curves. The mean ($\bar{b}_s$) and the observed variance ($V(b_s)$) of the standard slope were therefore calculated for each series

and are shown in Table 1. The predicted variance of b_s calculated from the average intra-assay statistics for each series (Table 1) was only two-thirds to one-half of the observed variance, a significant underestimate in most cases. In addition to the heterogeneity of the slopes, significant heterogeneity of assay residual variances (s^2), each with a large number of degrees of freedom, was found in each series by Bartlett's test.

DERIVATION OF M'S AND THEIR PREDICTED VARIANCES

M for each unknown was calculated on the basis of the combined slope (b_c) computed from all parallel responses in an assay. The predicted variance $V(M)$ was calculated by Bliss's formula, with Bliss's notation (3):

$$V(M) = \frac{s^2}{b_c^2}\left\{\frac{1}{N_s} + \frac{1}{N_u} + \frac{(\bar{y}_u - \bar{y}_s)^2}{B_c^2 - s^2t^2}\right\} \qquad (1)$$

The statistic $g = s^2t^2/B^2$ was calculated for a random sample of assays from each series using all parallel slopes in each assay. In most cases it was under .05 (Table 1), which is negligible by the usual criteria. The calculation of exact fiducial limits, using g, would thus be comparable to an expansion of $V(M)$ by about 5% at the most. Since this small change would have little effect on the analysis, it was ignored.

Replicate M's from Series I and II

One hundred sixty nine replicate M's for 75 urine extracts were available from ten assays in Series I and 52 assays in Series II. M and its variance estimated as described will be denoted as M_a and $V(M_a)$ for these replicates. In an attempt to assess the effect of the slope variability on the estimates, a second set of M's was computed, based on the mean standard slope ($\bar{b}_s$), as $M_z = \bar{x}_s - \bar{x}_u + (\bar{y}_u - \bar{y}_s)/\bar{b}_s$. Not unexpectedly, the individual M_z's differed considerably in some cases from their counterparts (M_a) based on assay slopes, but these differences were not systematic. A second set of variances $V(M_z)$ was also computed, in which $V(b_s)$, the observed standard slope variance, was included directly. The formula used was:

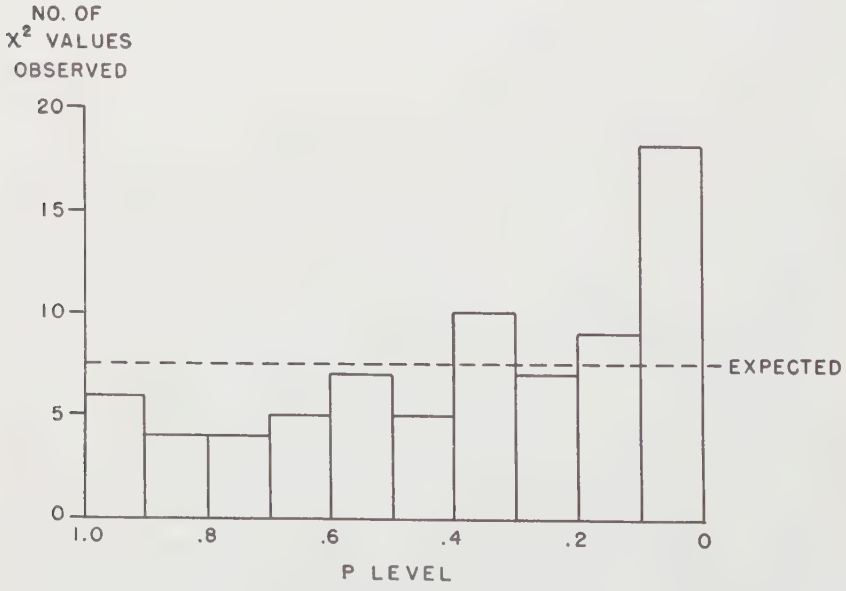
$$V(M_z) = \frac{1}{\bar{b}_s^2}\left\{s^2\left(\frac{1}{N_s} + \frac{1}{N_u}\right) + \frac{(\bar{y}_u - \bar{y}_s)^2 V(b_s)}{\bar{b}^2 - V(b_s)t_2^2}\right\} \qquad (2)$$

where all symbols have the meanings previously assigned, and t_2 is the tabular "t" value at P .05 for $V(b_s)$. The second term inside the curly brackets is equivalent to Bliss's $(\bar{y}_u - \bar{y}_s)^2 s_2^2/(B^2 - s_2^2 t_2^2)$ [3, p. 611].

The use of Equation 2 rests on two assumptions: that the mean difference $(\bar{y}_u - \bar{y}_s)$ is subject to the assay variance, and that M is also affected by the slope by assay interaction. The second assumption would necessitate modification of the model used by Finney [26] in his proof that M is unaffected by the slope variance, which is discussed below. The calculation of exact fiducial limits on the basis of these assumptions would be tantamount to an inflation of $V(M_x)$ by about 40% for M's from Series I and 26% for M's from Series II. As will be apparent from the data presented below, the adjustment would lessen but not completely remove the discrepancies between the observed and predicted error of M.

Replicate M's from Assay Series IV and V

From the assays in Series IV and V, replicate M's were available for two categories of unknowns: 56 duplicate estimates for 28 urine extracts, and 54 replicate estimates for 23 synthetic steroids (Table 2). They were derived from 21 assays in Series IV and 13 in Series V, but extracts and steroids were tested in different assays in the two series. M and $V(M)$ for these unknowns were computed with intra-assay statistics only.



DISTRIBUTION OF X^2 VALUES FOR ASSAY SLOPE M_a'S BY PROBABILITY (P) LEVEL

FIGURE 1

COMPARISON OF OBSERVED AND PREDICTED ERROR

The mean observed variance of the M's was computed from the sum of the squared deviations of replicate M's from their own mean. The mean observed and predicted variances and related statistics are compared in Table 3, for estimates from Series I and II, as computed both for M_a and for M_z. In both cases the difference between the observed and the predicted variance was of considerable magnitude and highly significant ($P < .001$), leading to the conclusion that, for these data, the predicted error was a gross underestimate of the actual error.

TABLE 3

VARIANCE OF 169 REPLICATED LOG POTENCY ESTIMATES (M) OF 75 UNKNOWNS FOR ANDROGEN ASSAYS IN SERIES I AND II, AS COMPUTED BY TWO METHODS

	Method of computation	
	Assay slopes (b_c)	Mean standard slope ($\bar{b}_s$)
Mean observed variance of M	0.0894	0.0769
Mean predicted variance of M	0.0281	0.0333
$F = \frac{\text{Mean observed variance}}{\text{Mean predicted variance}}$	3.18	2.31
$P(F)$	$< .001$	$< .001$
Percent standard error of R (antilog s.e.$_M$-1)100		
Observed	99	89
Predicted	47	52
Number of unknowns heterogeneous ($P < .05$)	13	12
Percent of unknowns heterogeneous ($P < .05$)	17.3	16.0

The observed and predicted variances were compared for each unknown individually by the X^2 test [3, 19]. Figure 1 shows the distribution of the probability values (P) corresponding to the values of X^2 obtained for the M_a's. The corresponding distribution (not shown) for the M_z's was similar, although the X^2's for individual unknowns were not identical for the two series of M's. The discrepancies from the expected distribution of probability values were highly significant ($P < .002$) in both sets of X^2, and the P .05 level was exceeded in more than 16% of the unknowns. The total distributions were characterized by a relative scarcity of low values and a surplus of high values.

Although the use of $\bar{b}_s$ and $V(b_s)$ in the computations narrowed the gap between the predicted and observed variances of M, the differences from the intra-assay results were not statistically significant. There was no convincing evidence that the use of $\bar{b}_s$ and $V(b_s)$ improved the accuracy of the predictions for individual unknowns.

In Series IV and V, the observed variance for urine extracts was almost exactly as predicted (Table 4). For synthetic steroids the predicted error of M was greater than for the urine extracts, mainly because the first M for a synthetic was usually based on a single dose. However, this relatively large predicted error significantly ($P < .05$) underestimated the observed variance in a ratio of 1:1.6.

TABLE 4

VARIANCE OF 110 REPLICATE M'S FOR 51 UNKNOWNS FROM ANDROGEN ASSAYS IN SERIES IV AND V

	Urine extracts	Synthetic steroids
Number of unknowns	28	23
Number of M's	56	54
Mean observed variance of M	0.0127	0.0277
Mean predicted variance of M	0.0121	0.0179
$F = \frac{\text{Mean observed variance}}{\text{Mean predicted variance}}$	1.05	1.55
$P(F)$	$> .05$	$< .05$
Percent standard error of R (antilog s.e.$_M$-1)100		
Observed	29.6	46.7
Predicted	28.8	36.1
Number of unknowns heterogeneous ($P < .05$)	2	3
Percent of unknowns heterogeneous ($P < .05$)	7.1	13.0

DISCUSSION OF THE VARIATION IN REPLICATE M'S

Although in the ethanol assays the predicted error of M was reduced, the marked improvement in inter-assay agreement could not have been anticipated from the intra-assay statistics (Table 1) or from comparisons among assays on a control chart. In the ethanol series, moreover, while the statistics provided accurate estimates of the precision of M for urine extracts, the predictions for synthetic steroids were inadequate. The findings strongly suggest that experimental error associated

with the administration of androgens in ether was the major factor behind the earlier results. However, the persistence of appreciable inter-assay error for steroid estimates indicates a need for continued investigation.

Systematic change with time was eliminated as a predominant cause of discrepancies by the failure to find any consistent trends. Cage differences may have contributed to the error since unknowns were not replicated over cages within an assay. However, several experiments with both ether and ethanol as solvents for the reference standard showed no significant cage differences, nor did duplicate cage means from 24 assays in Series IV (Table 5). It therefore seems unlikely that cage differences were an important source of error.

TABLE 5

ANALYSIS OF VARIANCE OF CODED MEAN RESPONSES TO ANDROSTERONE FROM 24 ASSAYS IN SERIES IV

In each assay, values for duplicate cage means of 11 chicks were assigned at random to "standard" or "unknown" group.

Source	D.f.	M.s.
Assays	23	133.1
Regression on dose	1	22,482.7***
Between "substances"	24	8.8
Assay × regression	23	19.1***
"Substance" × regression	24	11.8
Within cage	970	8.5
Mean M	$\bar{1}.9995$	
Observed variance of M	0.0137	
Predicted variance of M	0.0085–0.0339	

***$P < .001$

Lack of similarity in comb response to unknowns and to the standard androsterone may be a factor in the production of inter-assay error. Analysis of the data from all assays that met the previously outlined criteria for validity, did not produce suggestive evidence that non-identity or heterogeneity played an important role in the excess variation. However, further study of the possibility seemed indicated.

A related, more statistical issue considered was whether one could detect the presence of appreciable inter-assay error in M from direct assay analyses without such calculations on replicate estimates as were reported here. This would depend on identifying intra- or inter-assay

variation associated with variation in M, to provide a basis for estimating its variance as a more reliable indication of the actual error. The use of interactions of assays with treatments, slope and other effects as the error term would seem to have been ruled out by Finney [26] in the case of biological assays. He showed that in replicated analytical dilution assays differences in response levels are perfectly correlated with slope so that the within-group s^2 is the only suitable basis for estimating the error of M. The applicability of Finney's model to all biological assays has been questioned by Bliss [3, 27].

Experimental Trial

Information on some of these questions was sought in an experiment involving three substances: the standard androsterone (S), a chemically distinct androgen, methyl testosterone (MT), and an extract ($L25$) of a urine specimen obtained from a clinically normal young man. Since the predominant androgen in normal urines is androsterone it was expected that the responses to $L25$ and to S would be similar. Four dose levels of each substance were administered and replicated over four assays. Significant inter-assay slope variation and possibly additional interactions were anticipated. The main purposes were:

1) to see whether the three distinct substances would evoke different types of responses and whether MT would be different from the other two;

2) to test the applicability of Finney's model to this situation;

3) to estimate the intra- and inter-assay error in the ethanol assay method for each of the three substances and for comparisons among them; and

4) if appreciable inter-assay error were found, to seek statistically identifiable sources of variation associated with variation in M.

At least the usual opportunities for unidentified experimental error existed. In addition a recognized error in dosage made it necessary to discard one response group in assay 147. A "missing value" was computed in the usual manner [3] from the data of that assay for the combined analysis (Table 6). No significant deviations from linearity were found in any assay. The slope of the response to MT was non-parallel ($P < .05$) to the other slopes in all assays except number 147. In the combined analysis (Table 7) the significant findings of interest were the difference in the level of the response to MT and the variable differences between the regression for MT and the other regressions. The variance observed in M for each unknown was in agreement with prediction (Table 8).

TABLE 6

EXPERIMENT ON INTER-ASSAY ERROR
CODED RESPONSES*

Substance	Daily Dose	Assay 146	Assay 147	Assay 148	Assay 149
Androsterone (S)					
Dose: 1	0.1μg.	65.0	71.7**	74.5	80.6
2	0.2μg.	77.6	81.5	79.9	79.3
3	0.4μg.	87.6	90.2	92.1	88.8
4	0.8μg.	93.4	96.7	100.8	100.3
Mean		80.9	85.0	86.8	87.2
Slope		31.7	27.9	30.4	22.9
Methyl Testosterone (*MT*)					
Dose: 1	0.07μg.	65.7	68.8	66.7	67.7
2	0.14μg.	78.1	84.6	79.3	81.3
3	0.28μg.	80.2	90.0	90.1	90.8
4	0.56μg.	86.2	98.8	105.8	101.2
Mean		77.6	85.6	85.5	85.2
Slope		21.2	31.8	42.7	36.7
Extract (*L*25)					
Dose: 1	0.000025 d.	70.1	75.8	76.6	75.8
2	0.000050 d.	77.9	81.8	82.3	80.6
3	0.000100 d.	87.2	89.8	91.2	89.7
4	0.000200 d.	102.5	95.8	103.0	100.1
Mean		84.4	85.8	88.3	86.6
Slope		35.5	22.7	29.4	27.3
Mean slope		29.5	27.5	34.1	29.0
s^2 (per response mean of 11 chicks)		9.2	9.2	9.5	10.4

*Response metameter 100 log $\frac{\text{comb weight in mgm.}}{\text{body weight in decigrams}}$

**Missing value inserted.

All *L*25 doses in fractions of a 24-hour urine specimen.

The experiment provided gratifying evidence of the reduction in inter-assay error, but its potential usefulness in illuminating the problems was thereby reduced. The interpretation of the variable differences between *MT* and the other two substances is obscure. Although the non-parallelism would make *MT* formally non-assayable relative to

TABLE 7

EXPERIMENT ON INTER-ASSAY ERROR
COMBINED ANALYSIS OF VARIANCE ON CODED MEAN RESPONSES

Source of variation	D.f.	M.s.	
Assays	3	87.1***	
Substances	2	31.6*	
S vs. *L*25	1		12.8
S + *L*25 vs. *MT*	1		50.5*
Regression	1	4862.7***	
Substances × Regression	2	12.9	
S vs. *L*25	1		0.2
S + *L*25 vs. *MT*	1		12.3
Assays × Substances	6	9.4	
Assays × *S* vs. *L*25	3		6.2
Assays × *S* + *L*25 vs. *MT*	3		12.3
Assays × Regression	3	11.3	
Assays × Substances × Regression	6	22.8	
Assays × *S* vs. *L*25 × Regression	3		4.6
Assays × *S* + *L*25 vs. *MT* × Regression	3		41.0**
Deviations from regression	23	6.6	
Within groups	503	9.5	

*$P < .05$
**$P < .01$
***$P < .001$

TABLE 8

EXPERIMENT ON INTER-ASSAY ERROR
M VALUES FOR THE TWO UNKNOWNS

	Methyl Testosterone	Extract *L*25
Assay: 146	1.041	1.723
147	1.170	1.626
148	1.117	1.646
149	1.084	1.577
Unweighted mean	1.103	1.643
Mean observed variance	0.0030	0.0037
Mean predicted variance	0.0055	0.0054
$F = \frac{\text{Mean observed variance}}{\text{Mean predicted variance}}$	0.545	0.685
$P(F)$	$> .05$	$> .05$

androsterone, the agreement among replicate M values for MT was comparable to $L25$. A useful and apparently reliable estimate of its potency was therefore obtained despite the deviation from the analytical dilution model.

Except in the case of MT there was no evidence of appreciable inter-assay slope variance in the experiment, which therefore was not an adequate test of the applicability of Finney's model. However, the little evidence it provided tended to confirm the argument that assay interaction terms in the analysis of variance should not be the basis for predicting the error of M.

METHOD FOR INCORPORATING INTER-ASSAY ERROR INTO THE ESTIMATED PRECISION OF M

For estimates subject to inter-assay error, particularly the replicate M's from assay Series I and II, the best method for incorporating this error into individual estimates of precision required consideration.

Basic Hypothesis

The variance of the ith M for any unknown in these data ($i = 1, 2, \cdots$), derived from assay a, may be assumed to consist of $\sigma^2_{a_i} + \sigma^2_e$, where $\sigma^2_{a_i}$ is the intra-assay variance for that M_i, and σ^2_e is the inter-assay experimental error. The intra-assay variances ($\sigma^2_{a_i}$) were heterogeneous not only because of factors such as the number of dose levels administered, the number of chicks used, and the $\bar{y}_u$ obtained, but more fundamentally because of the inequality of the variance of individual responses (σ^2) in different assays. This assumption would place the M's under consideration in the category of estimates with heterogeneous intrinsic variances and with experimental error or interaction. Cochran [28] has shown that for such quantities, the weighted mean ($\bar{M}_w$) is an unsuitable combined estimate, and the choice lies between the unweighted mean ($\bar{M}$) and the generally more precise semiweighted mean ($\bar{M}_{sw}$). Cochran and Bliss [3], whose discussions were oriented mainly toward the combination of k estimates of one quantity when the available evidence consists of the data underlying these k estimates, have given formulas for the variances of $\bar{M}$ and $\bar{M}_{sw}$. Bliss has recommended that $\bar{M}_w$ be discarded in favor of $\bar{M}_{sw}$ when a X^2 test for heterogeneity is significant at P .05. However, the use of such a rule in analogous situations has been shown [29, 30] to result in underestimation of the variance in many cases, and more rigorous criteria have been considered [28, 30].

The effect of this rule was examined on the replicated potency estimates from assay Series I and II. The estimated variances of semi-

weighted means ($\bar{M}_{sw}$) for the 13 heterogeneous unknowns (Table 3) would be up to twenty times the predicted variance for $\bar{M}_w$. According to the same rule, $\bar{M}_w$ and its variance, including no correction for inter-assay error, would be computed for the 62 unknowns (83%) with non-heterogeneous potency estimates. If, as we have postulated, the M_i's for all 75 unknowns were actually samples from a distribution with a variance $\sigma^2_{ai} + \sigma^2_e$, the result would be an over-estimate of the error of apparently heterogeneous M's and an underestimate of the error of apparently homogeneous M's. The chance variation of individual M's around their true value would result in the assignment of vastly different degrees of precision to pooled estimates merely according to the value of P arbitrarily set as the criterion for the computation of $\bar{M}_{sw}$.

Limitations and Consequences of Incorporation of the Average Estimate ($\hat{\sigma}^2_e$) of Inter-Assay Error.

An alternative procedure for incorporation of inter-assay error into the estimated precision of M in a situation where replicates of multiple unknowns are available would involve the use of an average estimate, $\hat{\sigma}^2_e$, to be applied to all potency estimates. $\hat{\sigma}^2_e$ may be used both in the estimation of the variance of an M_i from a single assay, as $V(M_i) + \hat{\sigma}^2_e$, and in the calculation of semiweighted means and their variances. In our data, the mean observed variance of M_a, 0.0894, is equated to $\hat{\sigma}^2_{..} + \hat{\sigma}^2_e$ and $\hat{\sigma}^2_{..}$, as an estimate of the mean σ^2_{ai}, is equated to the mean $V(M_a)$, 0.0281. Then $\hat{\sigma}^2_e = 0.0894 - 0.0281 = 0.0613$. Thus, in the present case, $\hat{\sigma}^2_e$ is more than twice the mean $V(M_a)$, and its use would result, on the average, in the tripling of each estimate of variance. The limited precision thus assigned to M is a fair reflection of the assay error during the period when these estimates were obtained. The expanded variances of the M_i's would be employed in the usual manner for weighting the M_i's in the calculation of $\bar{M}_w$ and $V(\bar{M}_w)$. For any k estimates with n chicks per unknown in one assay, $V(\bar{M}_w)$ would be approximately $(\sigma^2_{ai} + \sigma^2_e)/k$, where σ^2_{ai} is a function of $1/n$. For any nk, there is an obvious advantage in increasing k as much as possible.

The reliability of the estimate of $\hat{\sigma}^2_e$ from our data is of course limited. The assays in the two series were represented unequally according to the number of M's from each one that met the criteria for inclusion in this study. Moreover, both the predicted and observed variances for the several M's from any one assay were obviously correlated. In spite of these reservations about the validity of a particular σ^2_e as a correction factor for inter-assay error, it would in our opinion provide a more accurate expression of the precision of M, whether replicated or not, than other methods which have been proposed.

DATA FROM ASSAY OF PARATHYROID ACTIVITY

An analysis of replicate M's from another parallel line assay method can be added to the above data. Twenty-four valid assays for the calcium mobilizing effect of parathyroid hormone [31] yielded 35 M values for 17 unknowns.* The residual error variance (s^2) and the assay slope (b_e) both showed considerable variation, and in a number of assays the slope was determined with poor precision, as indicated by the values for g (Table 9). The unknowns were extracts of parathyroid glands prepared by different methods. $V(M)$ calculated by Equation 1 showed considerable variation, but its mean value of 0.0352 compared well with an observed variance of 0.0307. The approximate X^2 for each unknown individually were significant at P .05 in two out of 17 instances (12%). As judged by these comparisons, the internal assay statistics in this method provided fairly reliable predictions of the observed error despite the variability of the assays.

TABLE 9

CHARACTERISTICS OF 35 REPLICATE M'S FOR 17 UNKNOWNS FROM 24 PARATHYROID ASSAYS

Mean assay slope ($\bar{b}_e$)	4.32
Mean assay residual variance (s^2)	1.33
Range of g values	0.06–0.76
Mean observed variance of M	0.0307
Mean predicted variance of M	0.0352

SUMMARY AND CONCLUSIONS

Replicate estimates (M) of the log androgenic potency of two types of unknowns, urine extracts and synthetic steroids, were made with a biological assay method of the parallel line type. During the first part of the study, in which the vehicle for administration of the androgenic substances was ether, the observed variance of 169 M's for 75 urine extracts was approximately three times as large as the variance predicted by the usual computations based exclusively on intra-assay error.

Rather than to divide the replicated M's from these assays arbitrarily into homogeneous and heterogeneous estimates and to treat them differently, it was considered preferable to treat all of them as samples

*Tabulations of the replicate M's are not included in this report, but mimeographed copies of these data are available and will be supplied on request.

from a distribution with a variance of $\sigma_{a_i}^2 + \sigma_e^2$, where $\sigma_{a_i}^2$ is the intra-assay variance and σ_e^2 is the inter-assay experimental error. The application of this concept to the computation of the variance of a single M_i or of pooled replicate M's was described.

In the second part of the study, in which the vehicle for administration of the androgens was changed from ether to ethanol, the observed variation between duplicate estimates for 28 urine extracts was in excellent agreement with the prediction. Thus a change in the assay procedure which permitted more accurate and reproducible administration resulted in the elimination of inter-assay error for urine extracts. However, in the same series of assays, the observed error of 54 replicate estimates for 23 synthetic steroids was still significantly greater (by 55%) than the predicted error.

The disappearance of inter-assay error for urine extracts and its persistence in the assay of synthetic steroids could not have been detected by an examination of the intra-assay statistics alone; it was necessary to calculate the variation among replicate estimates. This conclusion was also borne out by results of another bioassay method, for parathyroid hormone activity, in which 35 replicate M's for 17 unknowns, were examined. In spite of a great variability with respect to s^2 and slope and, in many cases, large values for g the variance of the replicated potency extimates was not significantly greater than the prediction from intra-assay statistics.

It is emphasized that although the total error of biological assays tends to approach the intra-assay error as a minimum, inter-assay error cannot safely be ignored until its absence has been demonstrated. Moreover, once the presence of a relatively large inter-assay error has been recognized and placed on a quantitative basis, it is possible to work more effectively for improvement of the true precision of a biological assay method by evaluating the effect of changes in procedure on the error between assays as well as within assays.

ACKNOWLEDGEMENT

We are indebted to Mrs. Elizabeth Moore, Miss Rita Nickerson, and Miss Katharine Hendrie for the statistical computations.

REFERENCES

[1] Sheps, M. C. and Munson, P. L. [1956]. The role of between-assay error in biological assays, *Biometrics, 12*, 534.

[2] Fieller, E. C. [1944]. A fundamental formula in the statistics of biological assay, and some applications, *Quart. J. Pharm. & Pharmacol. 17*: 117.

[3] Bliss, C. I. [1952]. *The Statistics of Bioassay*, Academic Press, New York.

[4] Bliss, C. I. [1946]. A revised cylinder-plate assay for penicillin, *J. Am. Pharm. A.* (Scient. Ed.) *35*: 6.

[5] Dews, P. B. and Berkson, J. [1954]. On the error of bio-assay with quantal response, in Kempthorne, O. et al.: *Statistics and Mathematics in Biology*, Iowa State College Press, Ames, Iowa, Chap. 27, p. 361.

[6] Humphrey, J. H., Mussett, M. V., and Perry, W. L. M. [1953]. The second international standard for penicillin, *Bull. Wld. Health Org. 9*: 15.

[7] Humphrey, J. H., Lightbown, J. W., Mussett, M. V., and Perry, W. L. M. [1953]. The international standard for aureomycin, *Bull. Wld. Health Org. 9*: 851.

[8] Humphrey, J. H., Lightbown, J. W., Mussett, M. V., and Perry, W. L. M. [1953]. The international standard for bacitracin, *Bull. Wld. Health Org. 9*: 861.

[9] Humphrey, J. H., Lightbown, J. W., Mussett, M. V., and Perry, W. L. M. [1954]. The international standard for dihydrostreptomycin, *Bull. Wld. Health Org. 10*: 901.

[10] Humphrey, J. H., Lightbown, J. W., Mussett, M. V., and Perry, W. L. M. [1955]. The international standard for oxytetracycline, *Bull. Wld. Health Org. 13*: 903.

[11] Miles, A. A., and Perry, W. L. M. [1953]. Biological potency and its relation to therapeutic efficacy, *Bull. Wld. Health Org. 9*: 1.

[12] Mussett, M. V., and Perry, W. L. M. [1955]. The international standard for thyrotrophin, *Bull. Wld. Health Org. 13*: 917.

[13] Mussett, M. V., and Perry, W. L. M. [1956]. The second international standard for corticotrophin, *Bull. Wld. Health Org. 14*: 543.

[14] Clarke, P. M. [1956]. A note on the combination of estimates of relative potency in multiple assays, *Biometrics 12*: 40.

[15] Allmark, M. G. [1951]. A collaborative study on the acute toxicity testing of several drugs, *J. Am. Pharm. A.* (Scient. Ed.) *40*: 27 .

[16] Nichols, A. B. [1953]. U. S. P. reference standards for biological assay, *J. Am. Pharm. A.* (Scient. Ed.) *42*: 215.

[17] Analytical Methods Committee [1954]. Report prepared by the thiamine (microbiological) panel of the sub-committee on vitamin estimations. The microbiological determination of thiamine, *Analyst 79*: 118.

[18] Jerne, N. K., and Wood, E. C. [1949]. The validity and meaning of the results of biological assays, *Biometrics 5*: 273.

[19] Finney, D. J. [1952]. *Statistical Method in Biological Assay*, Hafner, New York.

[20] Rakoff, A. E., Paschkis, K. E., and Cantarow, A. [1944]. Androgen assay on three-day-old male chicks, *Proc. Soc. Exper. Biol. and Med. 55*: 124.

[21] Munson, P. L., Goetz, F. C., Laidlaw, J. C., Harrison, J. H., and Thorn, G. W. [1954]. Effect of adrenocortical steroids on androgen excretion by adrenalectomized orchidectomized men, *J. Clin. Endocrinol. 14*: 495.

[22] Munson, P. L., and Sheps, M. C. [1956]. The biological assay of androgens, in Conference on Comparison of the Biological Properties of Steroids and Hormones, October 1954, Chicago, Committee on Research, A.M.A., p. 134.

[23] Harrison, J. H., Leman, C., Munson, P. L., and Laidlaw, J. C. [1955]. Hormone excretion before and after castration and adrenalectomy, *New England J. Med. 252*: 425.

[24] Harrison, J. H., Leman, C. B., Munson, P. L., and Laidlaw, J. C. [1955]. Androgen, gonadotropin, and steroid excretion in man before and after castration and adrenalectomy, *J. Urol. 73*: 580.

[25] Munson, P. L., Thorn, G. W., Harrison, J. H., and Sheps, M. C. [1955]. The effect of orchidectomy, ACTH administration, and adrenalectomy on the urinary excretion of androgens, determined by biological assay, Intern. Congr. Biochem. Abstr. of Commun., 3rd Congr., Brussels, 147.

[26] Finney, D. J. [1952]. On the precision of biological assays, *Acta Pharmacol. et Toxicol. 8*: 55.

[27] Bliss, C. I. [1952]. The clinical assay of diuretic agents. II. Estimation of the error in a clinical assay, *Biometrics 8*: 237.

[28] Cochran, W. G. [1954]. The combination of estimates from different experiments, *Biometrics 10*: 101.

[29] Bancroft, T. A. [1944]. On biases in estimation due to the use of preliminary tests of significance, *Ann. Math. Stat. 15*: 190.

[30] Paull, A. E. [1950]. On a preliminary test for pooling mean squares in the analysis of variance, *Ann. Math. Stat. 21*: 539.

[31] Munson, P. L. [1955]. Studies on the role of the parathyroids in calcium and phosphorus metabolism, *Ann. N. Y. Acad. Sci. 60*: 776.

INHERENTLY LOW PRECISION OF INFECTIVITY TITRATIONS USING A QUANTAL RESPONSE

G. G. Meynell

Department of Bacteriology, Postgraduate Medical School of London, Ducane Road, London, W.12, England

Infectivity* and pharmacological titrations appear at first sight to be very similar, not only in execution but also in interpretation. In both, the test preparation is serially diluted, a group of subjects is inoculated with each dilution and, in the case of quantal responses, to which this discussion is confined, the number of subjects responding to each dilution is recorded. The proportion of subjects responding is usually related to the logarithm of dose by a sigmoid curve which is roughly symmetrical. Finally, the potency of the preparation is often expressed in terms of the *ED*50, the dose causing 50% of subjects to respond, the precision with which this is done being clearly governed by the 'steepness' of the dose-response curve. The purpose of this paper is to suggest that the interpretations of infectivity and pharmacological titrations differ fundamentally owing to the completely different mechanisms by which the response is produced in the two cases and that infective particles act in a manner which restricts the steepness of the dose-response curve and consequently restricts the precision with which the *ED*50 of an infective preparation can be estimated. It will be shown that the greatest precision predicted for infectivity titrations is considerably less than that attainable in pharmacological titrations performed under the same conditions.

The production of an infective response is usually accounted for by a hypothesis derived from pharmacology which postulates that there exists an Individual Effective Dose (I.E.D.) for each subject such that a subject is certain to respond if it receives a dose equal to or greater than its I.E.D. The shape of the dose-response curve is considered to be governed solely by the distribution of I.E.D.s amongst the subjects (Finney [1952a], Wilson and Miles [1955]), assuming that errors in dosage can be ignored. This hypothesis offers a reasonable explanation for responses to particles, such as drug molecules, which are not self-reproducing. Such particles must cooperate to produce the response

*'Infective' is applied here to any system in which the response follows *multiplication* of the inoculated particles in the subject. Therefore it includes viable bacteria, viruses, tumour cells, protozoa, etc., and excludes drug molecules.

since, so far as is known, no response can be elicited by only one, or even a hundred, molecules. As even the smallest doses inoculated contain very large numbers of molecules, there will be negligible differences between replicate doses due to sampling error and variability in response must be due chiefly to variation in resistance of the subjects. The situation is quite different for infective particles which can multiply after administration. First, there are a few systems in which the subjects are completely susceptible and the inoculation of a single particle invariably causes a response. Such subjects are exactly similar to tubes of nutrient medium in bacterial counts performed by the dilution method and the shape of the dose-response curve is determined by the random distribution of particles amongst doses. Therefore, S, the proportion of subjects which does not respond to a mean dose of d viable particles is equal to e^{-d}, the first term of the Poisson series. This argument does not apply to the majority of infective systems for most subjects are only partially susceptible and the $ED50$ contains a considerable number of particles. Nevertheless, the hypothesis of the I.E.D. is not necessarily applicable to such systems for it is possible that the response results from the multiplication of only a small number, perhaps only one, of the many particles inoculated. This alternative hypothesis, apparently first stated by Halvorson [1935], has been referred to as the 'hypothesis of independent action' by Meynell and Stocker [1957]. It postulates that infective particles act completely independently after inoculation so that the fate of one does not affect the fate of another. Each viable particle then has a probability ($1 > p > 0$) of multiplying sufficiently to cause a response. If the subjects do not differ in resistance, S is again given by the first term of the Poisson series, e^{-pd}. A plot of S against log d gives a slightly asymmetrical sigmoid curve which closely resembles an integrated normal curve with standard deviation of 0.5 (Irwin [1942]). Thus, the hypothesis of independent action predicts that there will be considerable variability in response and a fairly flat dose-response curve, even if the subjects are all of the same resistance. The two hypotheses are incompatible, for the hypothesis of independent action implies that it is impossible to be completely certain that a dose of any size will cause a response since the outcome of challenge is assumed to be governed by chance. The hypothesis of independent action cannot yet be regarded as established and, as some workers may question the assumption of independence in systems where the particles are not of maximum virulence ($p < 1$), there follows a brief account of experimental evidence which suggests that this hypothesis is almost always valid.

The first two experiments are based on the following argument:

if independence obtains, only one particle will multiply sufficiently to produce a response in some of the subjects inoculated with many particles. This becomes very likely when a dose ≤ 1 *ED*50 is inoculated (Meynell and Stocker [1957] Fig. 2). Hence:

(a) In the first experiment, subjects are inoculated with a suspension containing equal proportions of distinguishable, equally virulent, variants of the same pathogen. Any subject which responds is sampled to determine the composition of the population of particles it then contains. If the hypothesis of independent action is valid, most samples from subjects responding to doses ≤ 1 *ED*50 should contain at least a predominance of one variant whereas samples from subjects responding to many *ED*50 should yield all the variants in the proportions present in the inoculum. The authors who first used this test of the hypothesis inoculated mixtures of different species and therefore did not test for interaction due to species-specific effects. Thus, Kunkel [1934] smeared leaves with a suspension containing equal numbers of *ED*50 of tobacco mosaic virus (T.M.V.) and aucuba mosaic virus (A.M.V.). Only a single dose was used (< 1 *ED*50) and each subject (an isolated lesion on the leaf surface in this system) contained only one or other of the two viruses. The experiment was repeated by Lauffer and Price [1945] who included doses > 1 *ED*50. These produced an unexpectably high proportion of lesions containing T.M.V. alone, presumably owing to antagonism (interference) between T.M.V. and A.M.V. when both were present in the same lesion. Liu and Henle [1953] inoculated eggs with $1/32 - 32$ *ED*50 of a mixture of influenza *A* and *B* viruses. Influenza *A* increases more rapidly than influenza *B* in this range of dosage with the result that eggs given 16 *ED*50 usually yielded a great excess influenza *A* after incubation. Nevertheless some eggs receiving doses < 8 *ED*50 yielded a great preponderance of influenza *B*, which strongly suggests that the fates of the inoculated particles were determined independently. A mixture of variants of the same species was first used by Zelle, Lincoln and Young [1946] who, using only one size of dose, exposed guinea-pigs to clouds containing four variants of *Bacillus anthracis*. Most of the fatally infected animals yielded one or other of two of the variants (which were presumably of higher virulence than the rest) but the results are difficult to interpret as neither the value of p (Druett, Henderson, Packman and Peacock [1953]) nor the mortality is known. Meynell and Stocker [1957] infected mice by intraperitoneal injection with a mixture of variants of either *Salmonella paratyphi B* ($p = 6.7 \times 10^{-6}$) or *Salmonella typhimurium* ($p \sim 10^{-3}$), the doses being in the range, $0.2 - 10^3$ *ED*50. As predicted, the proportion of the variants recovered varied greatly from mouse to mouse when the

dose was 1 *ED*50 or less, and became steadily more uniform (and similar to the inoculum) as the size of the dose increased. Nevertheless, far fewer mice than expected yielded only one variant. This discrepancy is attributed, on independent evidence, to a breakdown in resistance, caused by the outgrowth of the 'effective' fraction of the inoculum, which enabled members of the 'ineffective' fraction to multiply and to appear in the sample. Mice were also infected by mouth with two variants of *Salmonella typhimurium* (Meynell [1957]; $p \sim 5 \times 10^{-5}$) which were of equal virulence but unequal in their rates of increase in the subjects. Therefore, just as in the experiments of Liu and Henle [1953], nearly all mice inoculated with many *ED*50 yielded only one of the variants whereas those given 1 *ED*50 or less usually yielded either variant alone. The inoculation of a mixture of species or of variants of a single species has therefore always given results which are explicable on the hypothesis of independent action although the results may be disturbed by interaction (either antagonism or synergism) in the later stages of infection.

(b) In the second experiment, the relation of dose to the latent period intervening between inoculation and response is determined. In most systems, decrease in dosage prolongs the latent period (see, for example, Beard, Sharp and Eckert [1955]). However, if most responses to doses $\leq$ 1 *ED*50 are each due to the multiplication of a single particle, the latent period should tend to become constant for doses $\leq$ 1 *ED*50. This has been observed with mice given *Salmonella typhimurium* by intraperitoneal injection ($p = 10^{-3} - 10^{-6}$; Meynell and McCloy, in preparation). Dr. Bryan has kindly pointed out that plots of latent period against dose for several titrations of the Rous sarcoma virus (Bryan, Calnan and Moloney [1955]; Bryan [1956]) show apparently aberrant points for doses $\leq$ 1 *ED*50 which are compatible with this prediction.

In the third experiment, the effect produced by inoculation of a given number of particles in one dose is compared with the effect produced by the same total number of particles divided amongst doses which are inoculated either simultaneously at different sites or at different times by the same route. The hypothesis of independent action predicts that the proportion of subjects responding to a total dose of d particles will be the same whether or not the dose is divided. The predicted result has been obtained on the two occasions on which the prediction has been tested (Hewitt [1953], Goldberg, Watkins, Dolmatz and Schlamm [1954]).

Lastly, independence is a reasonable assumption to consider if it is borne in mind that the *LD*50 of a suspension of killed microorganisms

is far greater than the *LD*50 of live organisms. For example, the *LD*50 of dead Gram-negative bacteria (*Salmonella*, *Escherichia*) contains $c.10^{10}$ organisms while the *LD*50 of living bacteria is usually less than 10^7. Also, Maaloe [1948] and Rowley [1954] have shown that if the *LD*50 of an attenuated organism is measured with and without the addition of various numbers of killed bacteria of the same species, the size of the *LD*50 is unaffected unless more than 10^7 killed bacteria are included in the inoculum.

All the above points strongly suggest that the hypothesis of independent action is a useful model for most infections although interaction is known to occur in a few systems (Schneider and Zinder [1956], Gledhill [1956]).

The shape of the dose-response curve offers another means by which the validity of the hypothesis of independent action can be tested. As mentioned above, $S = e^{-pd}$, if the subjects are uniform in resistance. The dose-response curve will be flatter if the subjects differ in resistance (Armitage and Spicer [1956]) but its shape will not be altered by heterogeneity in the virulence of the particles (Fazekas de St. Groth and Moran [1955]). Hence, the observed relationship between dose and response can be compared with that predicted for uniform hosts, first, to see if the relationships could be the same (Druett [1952]) and, second, to see if the observed dose-response curve is significantly *steeper* than the predicted curve, a finding that would immediately show that the hypothesis of independent action was invalid. The value of p can be estimated from the data either by the methods given by Haldane [1939] and by Peto [1953], or by methods devised for bacterial counts by the dilution method (Finney, 1952b, §21.5). The observations can then be compared with the predictions in two ways: either by a χ^2 test (Haldane [1939]; or by the rapid test introduced by Moran [1954a, b]) to reveal discrepancies due to flattening of the observed curve. The latter test yields a quantity, M, a normal deviate, so that a value of $M \geq 1.645$ indicates a significant ($P \leq 0.05$) departure from expectation.

Table 1 gives details of infective systems in which the observed and predicted curves have been compared by at least one of these methods. Graphical comparisons can be found in Youden, Beale and Guthrie [1935], Bald [1937, 1950], Sprunt [1941], Parker, Bronson and Green [1941], Lauffer and Price [1945], Bang [1948], Kleckowski [1950], Fazekas de St. Groth and Cairns [1952], Goldberg *et al.* [1954], Beard, Sharp and Eckert [1955] and Eckert, Beard and Beard [1956]: all these curves appear compatible with or flatter than the predicted curve. The values of p given in the Tables indicate the maximum known values for each system. The values for viruses and tumours are only tentative

TABLE 1

DOSE-RESPONSE CURVES FROM INFECTIVITY TITRATIONS WHICH HAVE BEEN COMPARED WITH THE CURVE PREDICTED FOR UNIFORM HOSTS BY THE HYPOTHESIS OF INDEPENDENT ACTION.

Agent	Subject	Route of Administration	No. of titrations	p	P	M	b (95% confidence limits)	n	$\log_{10} f$	Author
					VIRUSES					
Myxoma virus	Rabbit	ID	1	—	<0.001	3.67	—	38–40	0.3	Parker [1940]
			1	—	—	1.09	—		0.3	P and M from Moran [1954, a]
Myxoma virus	Rabbit	ID	1	≤0.1	$0.1 > P > 0.05$	−2.04	—	20	0.3	Fenner & Woodroofe [1953]
Myxoma virus	Rabbit	ID	24	≤0.4	—	≤1.66	—	8	0.3	Fenner & McIntyre [1956]
Vaccinia virus	Rabbit	ID	4	? ≥ 0.25	<0.001–0.73	—	—	22–75	0.3	Parker [1938] P from Bryan and Beard [1940] p from Smadel, Rivers and Pickels [1939]
Vaccinia virus	Rabbit	ID	12	? ≥ 0.25	0.1–0.9	—	—	27–60	0.3	Sprunt and McDearman [1940]
Vaccinia virus	Rabbit	ID	1	$\leq 2 \times 10^{-3}$	0.67	—	—	?	0.3	Table I. Parker, Bronson & Green [1941]
Vaccinia virus	Rabbit	ID	6	—	0.03–0.92	—	—	10	0.3	Table III. Parker, Bronson & Green [1941]
Vaccinia virus	Rabbit	ID	6	—	0.001–0.48	—	—	10	0.3	Table IV. Parker, Bronson & Green [1941]
Influenza virus	Chick embryo	A	4	≥0.1	—	0.12–6.2	—	10	0.3	von Magnus [1951]
Influenza virus	Chick embryo	A	1	≥0.1	—	−1.53	—	10	0.3	p from Donald & Isaacs [1954]
Influenza virus	Chick embryo	A	1	≥0.1	—	1.9	—	10	0.6	p from Donald & Isaacs [1954]
Influenza virus	Chick embryo	A	1	≥0.1	—	0.69	—	50	1.0	p from Donald & Isaacs [1954]
Influenza virus	Chick embryo	A	146	≥0.1	—	0.21–2.04	—	5	0.5	Fazekas de St. Groth [1955]

TABLE 1—Continued

					VIRUSES					
Shope papilloma virus	Rabbit	ID	1	$\ll 1$	0.10	—	—	24	0.3	Bryan and Beard [1940]
Shope papilloma virus	Rabbit	ID	2	$\ll 1$	0.66	—	—	22	0.3	Bryan and Beard [1940]
Mouse encephalomyelitis virus	Mouse	IC	1	—	<0.001	—	—	—	—	Gard [1943]
Rickettsia tsutsugamushi	Mouse	IP	7	$\leq 10^{2} - 10^{-5}$	—	2.3–7.15	—	8–10	1.0	Fulton and Joyner [1945]
Eastern equine encephalomyelitis virus	Chick embryo	A	1	—	—	2.79	0.66 (0–1.4)	10	0.5	Crawley [1948]
Eastern equine encephalomyelitis virus	Chick embryo	A	1	—	—	0.48	1.82 (0.82–2.78)	9	0.5	Crawley [1948]
Newcastle Disease virus	Chick embryo	A	1	—	0.44	—	—	24–27	0.2	Nadel, Fryer and Eisenstarck [1957]
Newcastle Disease virus	Chick embryo	IC	1	—	0.009	—	—	36–42	0.3	Nadel, Fryer and Eisenstarck [1957]
Japanese encephalitis virus	Mouse	IC	3	—	<0.001	—	1.36	33–40	0.5	Matumoto, Arai and Iwasa [1950]
Plant viruses	Leaf	—	2	$\ll 1$	>0.05 – <0.05	—	—	c. 10^{3}	0.5	Kleckowski [1950] *p* from Knight [1950]
					BACTERIA					
Str. pneumoniae	Mouse	IP	1	1.0	>0.54		—	10	var.	Iwaszkiewicz and Neyman [1931] *p* from Morgan and Petrie [1933]
Salm. dublin	Mouse -normal	IP	1	0.17	0.28	—	—	17–20	0.3	Reid and MacLeod [1954]
Salm. dublin	-vaccinated	IP	1	9.7×10^{-3}	0.81	—	—	17–20	0.3	Reid and MacLeod [1954]
Salm. dublin	-vaccinated	IP	1	3×10^{-3}	0.03	—	—	17–20	0.3	Reid and MacLeod [1954]
Salm. dublin	-passively protected	IP	1	8.8×10^{-3}	0.38	—	—	17–20	0.3	Reid and MacLeod [1954]

TABLE I—Continued

Agent	Subject	Route of administration	No. of titrations	p	P	M	b (95% confidence limits)	n	$\log_{10} f$	Author
					BACTERIA					
Salm. typhimurium	Mouse	0	3	$\sim 10^{-5}$	—	2.1–3.6	0.7	7	1.0	Meynell [1957]
Salm. paratyphi B	Mouse	IP	3 pooled	6.7×10^{-7}	—	−0.91	1.81 (1.1–2.5)	15	0.5	Meynell and Stocker [1957]
Salm. typhi	Mouse	IP	8 pooled	2.3×10^{-8}	$0.5 > P > 0.1$	—	—	—	—	Peto [1953] p from Bacon, Burrows and Yates [1951]
Br. suis	Guinea pig	R	3 pooled	—	$0.9 > P > 0.5$	—	2.86 (1.67–4.05)	6–20	var.	Peto [1953] b, n & d from Elberg and Henderson [1948]
B. anthracis	Guinea pig	R	5 pooled	—	$0.5 > P > 0.1$	—	2.54 (1.69–3.39)	20–40	var.	Peto [1953] b, n & d from Druett *et al.* [1953]
					TUMOURS					
Krebs-2 ascites tumour	Mouse	IP	1	≥ 0.138	$0.3 > P > 0.2$	—	—	16–20	0.3	Hoskings, Meynell & Sanders [1956]
Krebs-2 ascites tumour	Mouse	IP	1	≥ 0.03	< 0.001	—	—	16–20	0.3	Hoskings, Meynell & Sanders [1956]

Route of inoculation: ID = intradermal; IC = intracerebral; IP = intraperitoneal; O = oral; R = respiratory; SC = subcutaneous; IV = intravenous; A = allantoic; YS = yolk sac.

Values of p marked $\leq$ are obtained by titration in subjects of differing resistance; those marked $\geq$ are obtained either by total particle counts or from physico-chemical data; and those marked $\sim$ come from titrations where the observed curve is not Poissonian.

The values of P are obtained from χ^2 tests.

M: Moran's statistic (Moran [1954, a, b]). b: slope estimated from the data by probit analysis. n: number of subjects per dose. f: dilution factor; var: varied.

as it is usually technically impossible to obtain an absolute estimate of d, the mean number of potentially infective (*i.e.* viable) particles inoculated (Hoskins, Meynell and Sanders [1956]; Isaacs [1957]). Relative estimates of p can be obtained in two ways. A suspension of particles can be titrated in two subjects of differing resistance, a method which yields a maximum value, as p may be less than unity for the more susceptible subject. Or, occasionally, the total number of particles is known either from direct counting or from chemical data: this provides a minimum value as an unknown fraction of the particles may be nonviable.

All save six of the curves summarised in Table 1 are compatible with or are flatter than the Poissonian curve, suggesting that the hypothesis of independent action is applicable to these systems. Five of the exceptional curves are from Table IV of Parker, Bronson and Green [1941]. The agent was an attenuated strain of vaccinia virus, producing very indistinct lesions, which was titrated separately in each of six rabbits. The authors considered that more than one lesion had to be present at each inoculation site to produce a visible response and the five exceptional curves are compatible with the curve predicted on the assumption that a visible response would be produced by ten or more lesions. The curve for the sixth rabbit was compatible with the predicted relationship $S = e^{-pd}$. The sixth exceptionally steep curve was reported by Nadel, Fryer and Eisenstarck [1957] who inoculated chicks with Newcastle Disease virus. This titration has apparently been performed only once and it would be of considerable interest to repeat it to establish that the incompatibility with prediction did not arise solely from sampling error which would be expected to cause an occasional discrepancy.

There are also many titrations reported which have not been compared with the Poissonian curve but have instead been analysed by probit methods. These can be compared with the predicted curve in a less exact manner. If $S = e^{-pd}$, as predicted for uniform subjects, probit S plotted against log d gives a slightly concave curve with slope, b, of 2.0003 at the $ED50$ point (Peto [1953]). This curve will approximate to a straight line with the same slope if the points are more or less symmetrically weighted around the $ED50$. Table 2 gives the slopes and 95% confidence limits for a number of infective systems. Only the titration with the greatest slope is given when more than one titration is reported. It will be seen that none of the slopes is significantly greater than 2, the approximate maximum value predicted by the hypothesis of independent action. Thus, the observations recorded in the Tables support the general validity of this hypothesis.

TABLE 2

DOSE-RESPONSE CURVES FROM INFECTIVITY TITRATIONS: SLOPES OBTAINED BY PROBIT ANALYSIS

Agent	Subject	Route	No. of titrations	p
				VIRUSES
Eastern equine encephalomyelitis virus	Chick embryo	A YS	49	—
Influenza virus	Tissue culture	—	3 pooled	$\leq 2.3 \times 10^{-2}$
Influenza virus	Chick embryo	A	2 pooled	≥ 0.1
Avian erythromyeloblastic leukosis virus	Chick	IV	3	$\sim 10^{-7}$
Rous sarcoma virus	Chicken	SC	1	≥ 0.02
Rous sarcoma virus	Chick	SC	2	≥ 0.02
				BACTERIA
Br. suis	Guinea pig	R	4	—
Br. suis	Guinea pig	R	7	—
Br. melitensis	Guinea pig	R	2	—
B. anthracis	Guinea pig	R	5	—
B. anthracis	Monkey	R	2	—
Salm. typhimurium	Mouse	IP	2 pooled	$\sim 10^{-3}$
Salm. typhi (+ mucin)	Mouse	IP	1	$\sim 8 \times 10^{-4}$
H. influenzae	Mouse	IC	>40	$\sim 10^{-3}$

The symbols are those used in Table 1.

CONCLUSIONS

1. The dose-response curve for infectivity titrations will always be relatively flat, even if subjects of uniform resistance could be used, and the slope obtained by probit analysis is unlikely to exceed 2. Infectivity titrations can therefore never be of the same precision, other

TABLE 2—Continued

Agent	b (95% confidence	n	$\log_{10} f$	Author
Eastern equine encephalomyelitis virus	2.02 (1.02 – 3.02)	9–11	0.5	Crawley [1948]
Influenza virus	1.74 (1.26–2.22)	55	1.0–0.7	Fulton and Armitage [1951]
Influenza virus	1.42 (0.97–1.86)	27	1.0–0.3	Fulton and Armitage [1951]
Avian erythromyeloblastic leukosis virus	0.379 (0.297–0.461)	115–119	0.7	Eckert, Beard and Beard [1951] p from Isaacs [1957]
Rous sarcoma virus	1.2	6–10	1.0	Bryan, Calnan and Moloney [1951] b from Fig. 4 p from Epstein [1956]
Rous sarcoma virus	1.81	20	1.0	Bryan [1956]
Br. suis	2.86 (1.67–4.05)	6–20	var.	Elberg and Henderson [1948]
Br. suis	2.38 (1.43–3.34)	20–50	var.	Druett, Henderson and Peacock [1956]
Br. melitensis	1.76 (1.00–2.51)	20	≤ 0.6	Elberg and Henderson [1948]
B. anthracis	2.54 (1.69–3.39)	20–40	var.	Druett, Henderson, Packman and Peacock [1953]
B. anthracis	3.19 (1.5–4.88	7–8	var.	Druett, Henderson, Packman and Peacock [1953]
Salm. typhimurium	0.66 (0.34–0.98)	12	0.48	Meynell and Stocker [1957]
Salm. typhi (+ mucin)	0.66	10	0.3–0.7	Batson [1949]
H. influenzae	1.46 (1.28–1.64)	15	0.6–0.7	Irwin and Standfast [1957], Table 23

The symbols are those used in Table 1.

things being equal, as pharmacological titrations, many of which have slopes in the range, 5–20 (Gaddum [1933], Bliss and Cattell [1943]).

2. The precision of infectivity titrations will not be greatly increased by the use of subjects specially bred for uniformity as variability in response does not result principally from heterogeneity in the resistance of the subjects. Moderate heterogeneity in resistance causes only a slight distortion of the curve predicted for uniform subjects (Armitage

and Spicer [1956]), which is presumably why many observed curves are compatible with the Poissonian curve.

3. If the observed and Poissonian curves are compatible, the concentration of infective particles can be rapidly estimated from the data by means of tables published for use with bacterial counts by the dilution method (Finney [1952b]).

SUMMARY

1. Experimental evidence suggests that infective (*i.e.* self-reproducing) particles act independently after inoculation, and do not cooperate as postulated by the hypothesis of the Individual Effective Dose.

2. Consequently, the dose-response curve for either completely susceptible subjects, or partially susceptible subjects of identical resistance, will be derived from the first term of the Poisson series when a quantal response is observed.

3. Considerable variability in response will therefore always be present in an infectivity titration using a quantal response, even if the subjects are of the same resistance, and this will be considerably greater than that observed in many pharmacological titrations. The observed variability will be increased if the subjects differ in resistance.

4. A search of the literature has shown that nearly all reported infectivity titrations yield dose-response curves which are compatible with these predictions.

I wish to thank Dr. P. Armitage, Dr. C. C. Spicer and Dr. B. A. D. Stocker for many valuable suggestions.

REFERENCES

Armitage, P. and Spicer, C. C. [1956]. The detection of variation in host susceptibility in dilution counting experiments. *J. Hyg. (Lond.), 54*, 401–414.

Bacon, G. A., Burrows, T. W. and Yates, M. [1951]. The effects of biochemical mutation on the virulence of *Bacterium typhosum:* the loss of virulence of certain mutants. *Brit. J. Exp. Path. 32*, 85–96.

Bald, J. G. [1937]. The use of numbers of infectives for comparing the concentration of plant virus suspensions. I. Dilution experiments with purified suspensions. *Ann. Appl. Biol. 24*, 33–55.

Bald, J. G. [1950]. Measurement of concentration of plant virus suspensions. Viruses 1950: California Institute of Technology, p. 17–29.

Bang, F. B. [1948]. Studies on Newcastle Disease Virus. I. An evaluation of the method of titration. *J. Exp. Med. 88*, 233–240.

Batson, H. C. [1949]. The relative significance of graded immunizing and challenge doses in measuring the potency of vaccines. *J. Exp. Med. 90*, 233–253.

Beard, J. W., Sharp, D. G. and Eckert, E. A. [1955]. Tumour viruses. *Advanc. Virus Res. 3*, 149–197.

Bliss, C. I. and Cattell, McK. [1943]. Biological assay. *Ann. Rev. Physiol. 5*, 479–539.

Bryan, W. R. [1956]. Biological studies on the Rous Sarcoma virus. IV. Interpretation of tumour-response data involving one inoculation site per chicken. *J. Nat. Cancer Inst. 16*, 843–863.

Bryan, W. R. and Beard, J. W. [1940]. Host influence in the characterization of response to the papilloma protein and to vaccinia virus. *J. Infect. Dis. 67*, 5–24.

Bryan, W. R., Calnan, D. and Moloney, J. B. [1955]. Biological studies on the Rous Sarcoma virus. III. The recovery of virus from experimental tumours in relation to initiating dose. *J. Nat. Cancer Inst. 16*, 317–335.

Crawley, J. F. [1948]. Statistical analysis of factors affecting the susceptibility of chick embryos to Eastern Equine Encephalomyelitis Virus. *J. Immunol. 59*, 83–94.

Donald, H. B. and Isaacs, A. [1954]. Counts of influenza virus particles. *J. Gen. Microbiol. 10*, 457–464.

Druett, H. A. [1952]. Bacterial invasion. *Nature (Lond.). 170*, 288.

Druett, H. A., Henderson, D. W., Packman, L. and Peacock, S. [1953]. Studies on respiratory infection. I. The influence of particle size on respiratory infection with anthrax spores. *J. Hyg. (Lond.). 51*, 359–371.

Druett, H. A., Henderson, D. W. and Peacock, S. [1956]. Studies on respiratory infection. III. Experiments with *Brucella suis*. *J. Hyg. (Lond.). 54*, 49–57.

Eckert, E. A., Beard, D. and Beard, J. W. [1951]. Dose-response relations in experimental transmissions of avian erythromyeloblastic leukosis. I. Host-response to the virus. *J. Nat. Cancer Inst. 12*, 447–463.

Eckert, E. A., Beard, D. and Beard, J. W. [1956]. Virus of avian erythroblastosis. I. Titration of infectivity. *J. Nat. Cancer Inst. 16*, 1099–1120.

Elberg, S. S. and Henderson, D. W. [1948]. Respiratory pathogenicity of Brucella. *J. Infect. Dis., 82*, 302–306.

Epstein, M. A. [1956]. The identification of the Rous virus; a morphological and biological study. *Brit. J. Cancer, 10*, 33–48.

Fazekas de St Groth, S. [1955]. Production of non-infective particles among influenza viruses: do changes in virulence accompany the von Magnus phenomenon? *J. Hyg. (Lond.). 53*, 276–296.

Fazekas de St Groth, S. and Cairns, H. J. F. [1952]. Quantitative aspects of influenza virus multiplication. IV. Definition of constants and general discussion. *J. Immunol. 69*, 173–181.

Fazekas de St Groth, S. and Moran, P. A. P. [1955]. A mathematical model of virus-cell interaction. Appendix to Fazekas de St Groth, S. [1955].

Fenner, F. and McIntyre, G. A. [1956]. Infectivity titrations of myxoma virus in the rabbit and the developing chick embryo. *J. Hyg. (Lond.). 54*, 246–257.

Fenner, F. and Woodroofe, G. M. [1953]. The pathogenesis of infectious myxomatosis: the mechanism of infection and the immunological response in the European Rabbit (Oryctolagus Cuniculus). *Brit. J. Exp. Path. 34*, 400–411.

Finney, D. J. [1952a]. *Probit Analysis*. 2nd ed, p. 8. Cambridge University Press, London.

Finney, D. J. [1952b]. *Statistical method in biological assay*. 1st ed., §21.5. Charles Griffin and Company, Ltd., London.

Fulton, F. and Armitage, P. [1951]. Surviving tissue suspensions for influenza virus titration. *J. Hyg. (Lond.). 49*, 247–262.

Fulton, F. and Joyner, L. [1945]. Cultivation of *Rickettsia tsutsugamushi* in lungs of rodents. *Lancet, 2*, 729–745.

Gaddum, J. H. [1933]. Reports on biological standards. III. Methods of biological assay depending on a quantal response. Spec. Rep. Ser. Med. Res. Coun., Lond., no. 183.

Gard, S. [1943]. Purification of poliomyelitis viruses. Experiments on murine and human strains. *Acta Med. Scand. Suppl. 143.*

Gledhill, A. W. [1956]. Quantitative aspects of the enhancing action of Eperythrozoa on the pathogenicity of Mouse Hepatitis Virus. *J. Gen. Microbiol. 15*, 292–304.

Goldberg, L. J., Watkins, H. M. S., Dolmatz, M. S. and Schlamm, N. A. [1954]. Studies on the experimental epidemiology of respiratory infections. VI. The relationship between dose of microorganisms and subsequent infection or death of a host. *J. Infect. Dis. 94*, 9–21.

Haldane, J. B. S. [1939]. Sampling errors in the determination of bacterial or virus density by the dilution method. *J. Hyg., (Lond.). 39*, 289–293.

Halvorson, H. O. [1935]. The effect of chance on the mortality of experimentally infected animals. *J. Bact. 30*, 330.

Hewitt, H. B. [1953]. Studies of the quantitative transplantation of mouse sarcoma. *Brit. J. Cancer, 7*, 367–383.

Hoskins, J. M., Meynell, G. G. and Sanders, F. K. [1956]. A comparison of methods for estimating the viable count of a suspension of tumour cells. *Exp. Cell Res. 11*, 297–305.

Irwin, J. O. [1942]. The distribution of the logarithm of survival times when the true law is exponential. *J. Hyg. (Lond.). 42*, 328–333.

Irwin, J. O. and Standfast, A. F. B. [1957]. The intracerebral mouse-protection test for pertussis vaccine. *J. Hyg. (Lond.). 55*, 50–80.

Isaacs, A. [1957]. Particle counts and infectivity titrations for animal viruses. *Advanc. Virus Res. 4*, 111–158.

Iwaszkiewicz, K. and Neyman, J. [1931]. Counting virulent bacteria and particles of virus. *Acta Biologiae Experimentalis, Varsovie. 6*, 101–142.

Kleczkowski, A. [1950]. Interpreting relationships between the concentrations of plant viruses and numbers of local lesions. *J. Gen. Microbiol. 4*, 53–69.

Knight, C. A. [1950]. Some salient points concerning plant viruses. Viruses 1950: California Institute of Technology, p. 52–54.

Kunkel, L. O. [1934]. Tobacco and aucuba-mosaic infections by single units of virus. *Phytopathology, 24*, 13.

Lauffer, M. A. and Price, W. C. [1945]. Infection by viruses. *Arch. Biochem. 8*, 449–468.

Liu, O. C. and Henle, W. [1953]. Studies on host-virus interactions in the chick embryo-influenza virus system. VII. Data concerning the significance of infectivity titration end-points and the separation of clones at limiting dilutions. *J. Exp. Med. 97*, 889–902.

Maaløe, O. [1948]. I. Pathogenic-apathogenic transformation of *Salmonella typhimurium*. *Acta Path. Microbiol. Scand. 25*, 414–430.

Magnus, P. von. [1951]. Propagation of the PR8 strain of influenza A virus in Chick embryo. I. The influence of various experimental conditions on virus multiplication. *Acta Path. Microbiol. Scand. 28*, 250–277.

Matumoto, M., Arai, H. and Iwasa, S. [1950]. Theorical study on dosage-response curve of viruses. *Jap. J. Exp. Med. 20*, 803–816.

Meynell, G. G. [1957]. The applicability of the hypothesis of independent action to fatal infections in mice given *Salmone'la typhimurium* by mouth. *J. Gen. Microbiol. 16*, 396–404.

Meynell, G. G. and Stocker, B. A. D. [1957]. Some hypotheses on the aetiology of

fatal infections in partially resistant hosts and their application to mice challenged with *Salmonella paratyphi-B* or *Salmonella typhimurium* by intraperitoneal injection. *J. Gen. Microbiol. 16*, 38–58.

Moran, P. A. P. [1954a]. The dilution assay of viruses. *J. Hyg., (Lond.). 52*, 189–193.

Moran, P. A. P. [1954b]. The dilution assay of viruses. II. *J. Hyg., (Lond.). 52*, 444–446.

Morgan, W. T. J. and Petrie, G. F. [1933]. The titration of anti-pneumococcus serum (Type I) in mice. *Brit. J. Exp. Path. 14*, 323–338.

Nadel, M. K., Fryer, H. C. and Eisenstarck, A. [1957]. The probable minimum numbers of Newcastle Disease virus particles required to initiate infection in chick embryos and baby chicks. *J. Infect. Dis. 100*, 88–91.

Parker, R. F. [1938]. Statistical studies on the nature of the infectious unit of vaccine virus. *J. Exp. Med. 67*, 725–738.

Parker, R. F. [1940]. Studies of the infectious unit of myxoma. *J. Exp. Med. 71*, 439–444.

Parker, R. F., Bronson, L. H. and Green, R. H. [1941]. Further studies of the infectious unit of vaccinia. *J. Exp. Med. 74*, 263–281.

Peto, S. [1953]. A dose-response equation for the invasion of micro-organisms. *Biometrics, 9*, 320–335.

Reid, D. B. W. and MacLeod, D. R. E. [1954]. The relation between dose and mortality of *Salmonella dublin*. *J. Hyg. (Lond.). 52*, 18–23.

Rowley, D. [1954]. The virulence of strains of *Bacterium coli* for mice. *Brit. J. Exp. Path. 35*, 528–538.

Schneider, H. A. and Zinder, N. D. [1956]. Nutrition of the host and natural resistance to infection. V. An improved assay employing genetic markers in the double strain inoculation test. *J. Exp. Med. 103*, 207–223.

Smadel, J. E., Rivers, T. M. and Pickels, E. G. [1939]. Estimation of the purity of preparations of elementary bodies of vaccinia. *J. Exp. Med. 70*, 379–385.

Sprunt, D. H. [1941]. The effect of the virus: host cell relationship on infection with vaccinia. *J. Exp. Med. 74*, 81–89.

Sprunt, D. H. and McDearman, S. [1940]. Studies on the relationship of sex hormones to infection. III. A quantitative study of the increased resistance to vaccinial infection produced by the estrogenic hormone and pseudopregnancy. *J. Immunol. 38*, 81–95.

Wilson, G. S. and Miles, A. A. [1955]. *Topley and Wilson's Principles of Bacteriology and Immunity*. 4th. ed. Arnold, London p. 1132.

Youden, W. J., Beale, H. P. and Guthrie, J. D. [1935]. Relation of virus concentration to the numbers of lesions produced. Contr. Boyce Thompson Inst. *7*, 37–53.

Zelle, M. R., Lincoln, R. E. and Young, G. A. [1946]. Respiratory pathogenicity of *Bacillus anthracis* spores. II. Genetic variation in respiratory pathogenicity and invasiveness of colonial variants of *B. anthracis*. *J. Infect. Dis. 79*, 247–253.

MULTIPLE RANGE TESTS FOR CORRELATED AND HETEROSCEDASTIC MEANS*

DAVID B. DUNCAN
Universities of Florida and North Carolina, U.S.A.

1. INTRODUCTION

Multiple range tests have been developed by several writers, for example, D. Newman [8], M. Keuls [5], J. W. Tukey [10] and D. B. Duncan [3], for testing differences between several treatment means in cases in which all such differences are of equal *a priori* interest. These tests, which are also described in recent textbooks, for example, W. T. Federer [4, chapter 2], have been worked out for data in which the treatment means are homoscedastic (have equal variances) and are uncorrelated. Recently, C. Y. Kramer [6] has presented a simple method for extending these procedures to give useful tests for differences between means with unequal replications, the method being applicable to any set of heteroscedastic uncorrelated means. In a subsequent paper [7], the same author has given further extensions to tests of means which are also correlated, such as the adjusted means from analyses of covariance or from incomplete block designs. Similar work has also been done by E. Bleicher [1] and P. G. Sanders [9] in extending a multiple F test to making tests in lattice and rectangular lattice designs.

One purpose of this paper is to present a more complete method for these extensions which necessarily sacrifices a little in simplicity but is more powerful, especially in cases in which the differences between the means have appreciably different variances. Another purpose is to indicate briefly the closeness of the properties of these complete tests of heteroscedastic and correlated means to those of the corresponding tests of homoscedastic and uncorrelated means. Incidental to these main purposes, a short-cut skipping principle, useful in applying multiple range tests to a large number of treatment means (or totals), is also presented.

*Research jointly supported by the Florida Agricultural Experiment Station, by the U. S. Public Health Service and by the U. S. Air Force through the Office of Scientific Research of the Air Research and Development Command.

2. BASIC RULE FOR COMPLETE TESTS

Let m_1 , m_2 , $\cdots$, m_n represent n normally distributed means such that the variance of the difference between each pair can be written $V(m_i - m_j) = k_{ij}\sigma^2$ where k_{ij} is known and σ^2 is an expected error mean square. Let s^2 with n_2 degrees of freedom be the usual type of analysis of variance error mean square estimate for σ^2. In other words, $n_2 s^2/\sigma^2$ is distributed as $\chi^2_{n_2}$ and is independent of the means m_1 , m_2 , $\cdots$, m_n .

Call $a_{ij} = \sqrt{2/k_{ij}}$ the *adjustment factor* for and $(m_i - m_j)' = a_{ij}(m_i - m_j)$ the *adjusted difference* between the means m_i and m_j , and call $R'_p = s \cdot z_p$ the *critical value* for p means, where z_p is the Studentized significant range for p means for n_2 degrees of freedom and for an α-level test.

The proposed complete basic rule for an α-level multiple range test may then be expressed as follows: ANY SUBSET OF p MEANS IS HOMOGENEOUS IF THE LARGEST ADJUSTED DIFFERENCE IN THE SUBSET FAILS TO EXCEED THE CRITICAL VALUE R'_p . ANY TWO MEANS NOT BOTH CONTAINED IN THE SAME HOMOGENEOUS SUBSET ARE SIGNIFICANTLY DIFFERENT. ANY TWO MEANS BOTH CONTAINED IN THE SAME HOMOGENEOUS SUBSET ARE NOT SIGNIFICANTLY DIFFERENT.

This is the same as the basic rule implicitly adopted by Kramer [6] except that in the latter a subset of p means is declared homogeneous if its *adjusted range* does not exceed R'_p . If an adjusted difference within a subset exceeds the adjusted range, as it may do through having a smaller variance and hence a larger adjustment factor, it will be significant by the complete rule and this may also result in the detection of further significant differences.

3. NUMERICAL EXAMPLE I: TEST OF UNEQUALLY-REPLICATED MEANS

Table 1 illustrates a convenient method for applying the complete rule. The example consists of the application of a 5 per cent level new multiple range test [3] to a set of seven unequally-replicated treatment means from a completely randomized design. A similar extension of any multiple range test, e.g., [5], [8], and [10], could be made by the same method, the only difference being in the source used for the Studentized significant ranges z_p , in section (b). Table 2 gives details of the computation of the adjusted differences used in Table 1.

The initial preparation of the data is the same as for Kramer's method [6]. Table 1, section (a) shows the analysis of variance concluding with the calculation of the error standard deviation $s = 73.45$. Section (b) shows the computation of the critical values $R'_p = s \cdot z_p$, the Studentized significant ranges z_p having been taken from [3, Table II]

TABLE 1

5 PER CENT LEVEL NEW MULTIPLE RANGE TEST [3] OF SEVEN UNEQUALLY-REPLICATED MEANS

a) *Analysis of Variance*

Source	*D.f.*	*M.s.*	$s = \sqrt{m.s.}$
Between treatments	6		
Error	16	5,394.6	73.45

b) *Critical Values:* $R'_p = s.z_p$

p:	(2)	(3)	(4)	(5)	(6)	(7)
z_p:	3.00	3.15	3.23	3.30	3.34	3.37
R'_p:	220.4	231.4	237.2	242.4	245.3	247.5

c) *Ranked Treatment Means and Replication Numbers*

D	*F*	*A*	*B*	*C*	*E*	*G*
680	734	743	851	873	902	945
(3)	(2)	(5)	(5)	(3)	(2)	(3)

d) *Test Sequences*

Seq.	*Steps*	*Result*
1.	$(G-D)' > R'_7$, $(G-F)' > R'_6$, $(G-A)' > R'_5$, $(G-B)' \not> R'_4$.	$(BCEG)$
2.	$(E-D)' > R'_6$, $(E-F)' \not> R'_5$. $FABCE$: $(E-A)' > R'_5$.	
	$FBCE$: $(E-F)' > R'_4$. BCE.	—
3.	$(C-D)' > R'_5$, $(C-F)' \not> R'_4$. $FABC$: $(C-A)' > R'_4$.	
	FBC: $(C-F)' \not> R'_3$, $(C-B)' \not> R'_3$, $(B-F)' \not> R'_3$.	(FBC)
4.	$(B-D)' > R'_4$, $(B-F)' \not> R'_3$. FAB: $(B-A)' > R'_3$. FB.	—
5.	$(A-D)' \not> R_3$.	(DFA)

e) *Final Result*

(DFA) (FBC) $(BCEG)$

Any two means *not appearing together* within the same parentheses are *significantly different.* Any two means *appearing together* within the same parentheses are *not significantly different.*

for a 5 per cent level test entering at the row for $n_2 = 16$ degrees of freedom. Section (c) shows the treatment means ranked in ascending order together with their respective replication numbers in parentheses. In any test of uncorrelated means it is helpful to list under the ranked means measures, such as replication numbers in this case, which provide a quick method of visually assessing the relative magnitudes of the variances of the means and hence of the variances of the differences between them.

The main part of the test is in the sequences of steps in section (d). Each step consists of an application of the basic rule to a particular subset. Sequence 1 consists of steps involving all subsets in which

the top mean G is the largest mean, sequence 2 involves all subsets in which the second mean E is the largest mean, and in general, sequence i involves all subsets in which the i-th mean is the largest mean.

The order of steps in each sequence is the same at the beginning as that of previous procedures [3] and [6]. The steps in sequence 1, for example, consist of testing the adjusted ranges first of the whole set $DFABCEG$, then of the subset $FABCEG$, then of $ABCEG$ and so on. At each step the lowest mean is dropped to give the subset for the next test.

TABLE 2

ARITHMETICAL DETAILS FOR CALCULATING ADJUSTED DIFFERENCES IN TABLE 1, SECTION (d)

$$a_{ij} = \sqrt{2/k_{ij}} = \sqrt{2\Big/\left(\frac{1}{r_i} + \frac{1}{r_j}\right)} = \sqrt{2r_ir_j/(r_i + r_j)},$$

where $r_1, r_2, \cdots, r_7$ are the replication numbers for the respective means, thus

$$(G - D)' = (G - D)a_{GD} = 265\sqrt{2(3)(3)/6} = 265(1.732) = 459.0.$$

Similarly

$(G - F)' = 211(1.549) = 326.8,$ $(G - A)' = 202(1.936) = 391.1,$
$(G - B)' = 94(1.936) = 182.0,$ $(E - D)' = 222(1.549) = 343.9,$
$(E - F)' = 168(1.414) = 237.6,$ $(E - A)' = 159(1.690) = 268.7,$
$(C - D)' = 193(1.732) = 334.3,$ $(C - F)' = 139(1.549) = 215.3,$
$(C - A)' = 130(1.936) = 251.7,$ $(C - B)' = 22(1.936) = 42.6,$
$(B - F)' = 117(1.690) = 197.7,$ $(B - D)' = 171(1.936) = 331.1,$
$(B - A)' = 108(2.236) = 241.5,$ $(A - D)' = 63(1.936) = 122.0.$

The changes in the complete test come in each sequence when the adjusted range of a subset of p means fails to exceed R'_p. If the adjustment factor for the range is smaller than that of any other difference in the subset, or, in a test like this of unequally-replicated means, if either of the extreme means has fewer replications than any of the other means, any of the other differences with a larger adjustment factor, should also be tested. In such cases it is helpful to write down the subset concerned as a reminder that it will still be the subset under test until an adjusted difference within it is found to exceed R'_p. For example, when the adjusted range $(C - F)'$ of $FABC$ fails to exceed R'_4 in step 2 of sequence 3, the subset $FABC$ is written down before testing the adjusted difference $(C - A)'$ in the next step. This serves as a reminder that $(C - A)'$ must be compared with R'_4 and not R'_3 as would otherwise

have happened. Similarly in the next three steps, the preliminary recording of FBC serves as a reminder that $(C - F)'$, $(C - B)'$ and $(B - F)'$ each have to be compared with R'_3.

When an adjusted difference between the top mean in a subset of p means and an internal mean is found to exceed R'_p, the internal mean is dropped to give the next subset to be tested. For example, in step 3 of sequence 3 $FABC$ is reduced to FBC in this way by dropping A when $(C - A)'$ exceeds R'_4.

When an adjusted difference, not involving the top mean, is found to exceed R'_p, two subsets may qualify for further testing in the same sequence. For example, if four means were ranked and had replication numbers as follows

P	S	R	Q
(1)	(50)	(100)	(1)

the testing steps could be

$$(Q - P)' \not> R'_4 \,.\, PSRQ\text{: } (Q - S)' \not> R'_4 \,,\, (Q - R)' \not> R'_4 \,,$$

$$(R - P)' \not> R'_4 \,,\, (R - S)' > R_4 \,,$$

and the subsets PRQ and PSQ would both qualify for testing in further steps in the same sequence.

The testing for a subset terminates either when it is shown to be homogeneous, which fact is recorded by noting the subset in parentheses in the result column at the end of the sequence involved, or when the subset is found to be completely included within another subset already shown to be homogeneous. For example, $BCEG$ is recorded $(BCEG)$ at the end of the first sequence to denote its homogeneity. This follows from the fact that the adjusted range $(B - G)'$ of $BCEG$ does not exceed R'_4 and neither of the replications 5 and 3 of B and G is less than the replications 3 or 2 of C or E. The result (DFA) of sequence 5 is of a similar form. In other cases, e.g. the result (FBC) in sequence 3, it is sometimes necessary to test each adjusted difference in the subset before it can be declared homogeneous.

Sequences 2 and 4 provide no additional homogeneous subsets because they terminate at subsets BCE and FB, which are included in $(BCEG)$ and (FBC) respectively already shown to be homogeneous.

It should be noted in conclusion that it is possible for more than one homogeneous subset to be found in a single sequence. For example, in the case of the means $PSRQ$ discussed above the sequence involved could terminate with the results (PRQ) and (PSQ) or even (PQ) (SQ) and (RQ) depending on the other data involved.

Section (c) of Table 1 shows a useful way of presenting the results of

the test. The device of presenting the whole set with homogeneous subsets underscored as is done in [3] and [6] cannot be used here because of the differences in the order of the means in the various homogeneous subsets. For example, A is to the right of F in (DFA) but not in (FBC). The new method of putting homogeneous groups in parentheses can also be used in tests of equally replicated means and may be preferred for printing purposes.

4. NUMERICAL EXAMPLE II: TEST OF TREATMENT TOTALS IN A SIMPLE LATTICE DESIGN (Including the Use of Skipping Short Cuts).

Table 3 illustrates the application of a similar 5 per cent level test to the adjusted totals in a 5×5 simple lattice design. The data are those given by Cochran and Cox [2, section 10.29] for a design with two repetitions. Table 4 gives additional details of the computation of the adjusted differences used in Table 3.

Section (a), Table 3, shows the value of s obtained from the error mean square s^2 (denoted E_c in [2]) for the experiment and its degrees of freedom n_2. Section (b) shows the adjustment factors for differences between treatment totals (totals being more convenient than means to use in a case like this).

Cochran and Cox give $2E_c[1 + (n - 1)\mu]/r$, (their n being the number (2) of repetitions involved) for the estimated variance of a difference between two means for treatments in the same block. Thus, using $k_{++}\sigma^2$ to denote the variance of a difference between totals for treatments in the same block we have $k_{++} = 2r[1 + (n - 1)\mu]$. Then using a_{++} for the corresponding adjustment factor, we have $a_{++} = \sqrt{2/k_{++}} = (r[1 + (n - 1)\mu])^{-(1/2)}$. Similarly, if a_{+-} is used to denote the adjustment factor for differences between totals of treatments not in the same block, we have $a_{+-} = \sqrt{2/k_{+-}} = (r[1 + n\mu])^{-(1/2)}$. In this example, $r = 4$, $n = 2$, $\mu = 0.1270$ and the adjustment factors work out to be as shown in section (b).

Section (c) shows the ranked treatment totals and critical values required for the test. The arrangement is different from the corresponding sections of the previous example solely because of the largeness of the number (25) of treatments involved. The new arrangement is convenient for applying a skipping method which short cuts many of the steps at the beginning of each sequence. In all other respects the procedure is unchanged. In the first column the treatments $1, 2, \cdots, 25$ as they have been denoted in [2] are redenoted $A, B, \cdots, Y$ for convenience in the recording of treatment subsets. The number $(i \cdot j)$ appearing after each treatment letter denotes the blocks in which the

treatment falls. Thus (3.1) after treatment K shows that it belongs to block 3 and to block 1 in the first and second types of replicates, respectively. These numbers are useful in indicating which treatments do and which do not appear together in the same block and thus which adjustment factor applies to each difference.

The second column of section (c) shows the adjusted treatment totals from [2] followed by doubly adjusted treatment totals in parentheses which, for brevity, we will call treatment totals and adjusted

TABLE 3.

5 Per Cent Level Multiple Range Test of Adjusted Treatment Totals from a 5 × 5 Simple Lattice Design

a) *From Analysis of Variance*

$n_2 = 56, \qquad s^2 = 13.60, \qquad s = 3.69$

b) *Adjustment Factors for Differences between Treatment Totals*

Two treatments in same block: $a_{++} = .471$
Two treatments not in same block: $a_{+-} = .447$

c) *Ranked Treatment Totals and Critical Values*

Treatment	*Total*	p	z_p	R'_p
11K(3.1)	88.4(39.5)	25	3.47	12.80
2B(1.2)	77.3(34.6)	24	3.47	12.80
15O(3.5)	74.7(33.4)	23	3.47	12.80
14N(3.4)	71.6(32.0)	22	3.47	12.80
24X(5.4)	70.6(31.6)	21	3.47	12.80
22V(5.2)	68.1(30.4)	20	3.47	12.80
1A(1.1)	66.6(29.8)	19	3.46	12.77
21U(5.1)	61.4(27.4)	18	3.45	12.73
4D(1.4)	58.8(26.3)	17	3.44	12.69
16P(4.1)	58.3(26.1)	16	3.43	12.66
23W(5.3)	55.7(24.9)	15		
25Y(5.5)	52.7(23.6)	14		
13M(3.3)	52.7(23.6)	13		
18R(4.3)	52.6(23.5)	12		
20T(4.5)	51.6(23.1)	11		
12L(3.2)	51.1(22.8)	10		
5E(1.5)	50.9(22.8)	9		
7G(2.2)	47.6(21.3)	8	3.28	12.10
6F(2.1)	46.9(21.0)	7	3.25	11.99
10J(2.5)	46.2(20.7)	6		
17Q(4.2)	46.0(20.6)	5		
8H(2.3)	45.2(20.2)	4		
3C(1.3)	44.9(20.1)	3		
9I(2.4)	38.1(17.0)	2		
19S(4.4)	21.5(9.6)	—		

TABLE 3—(Continued)

d) *Test Sequences*

Seq.	*Steps*	*Result*
1.	$39.5 - R'_{25} = 26.70$, $39.5 - R'_{8} = 27.4$.	
	$a_{++}(K - U) > R'_{8}$, $a_{++}(K - A) \not> R'_{7}$.	$(AVXNOBK)$
2.	$34.6 - R'_{24} = 21.80$, $34.6 - R'_{16} = 21.94$.	
	$a_{++}(B - E) \not> R'_{16}$.	$(ELTRMYWPDUAVXNOB)$
3.	$33.4 - R'_{23} = 20.60$, $33.4 - R'_{19} =$	
	20.63, $33.4 - R'_{18} = 20.67$.	
	$a_{++}(O - J) > R'_{18}$. $FGEL \cdots O$: —	$(FGELTRMYWPDUAVXNO)$
4.	$32.0 - R'_{22} = 19.20$, $32.0 - R'_{20} = 19.20$	
	$CHQJ \cdots N$:—	$(CHQJFGELTRMYWPDUAVXN)$
5.	$31.6 - R'_{21} = 18.80$.	—
6.	$30.4 - R'_{20} = 17.60$.	—
7.	$29.8 - R'_{19} = 17.03$.	—
8.	$27.4 - R'_{18} = 14.67$, —	$(ICHQJFGELTRMYWPDU)$
last	$9.6 + R'_{17} = 22.29$, $9.6 + R'_{8} = 21.70$,	
	$SICHQJFG$: $a_{++}(Q - S) \not> R'_{8}$.	$(SICHQJFG)$.

e) Final Results

$(SICHQJFG)(ELTRMYWPDUAVXNOB)$
$(ICHOJFGELTRMYWPDU)(AVXNOBK)$
$(CHOJFGELTRMYWPDUAVXN)$
$(FGELTRMYWPDUAVXNO)$

Any two treatments *not appearing together* within the same parentheses are *significantly different.* Any two treatments *appearing together* within the same parentheses are *not significantly different.*

treatment totals, respectively. Each adjusted treatment total in parentheses is obtained by multiplying the corresponding treatment total by the smallest adjustment factor. In this example there are only two adjustment factors, the smaller being .447, hence the adjusted total for K, for instance, is $88.4(.447) = 39.5$ as shown. The column of adjusted totals is a new feature needed in the skipping short cut steps.

The last two columns of section (*c*) show the Studentized ranges z_p and the critical values $R'_p = 3.69z_p$. The middle column for p helps in identifying the z_p and R'_p values. The z_p values in this example are obtained from [3, Table II] for a 5 per cent level test the same as in the previous example except that for larger values of p some simple linear interpolation is needed. When a large number of treatments is involved as in this example, not all of the critical values R'_p will be required. Each one should thus be obtained only as needed in the sequence steps.

In Table 3, for instance, only 12 of the possible 24 R'_p values are ultimately found to be needed.

Section (d) shows the main part of the test arranged in steps within sequences as in the previous example. The first two steps in sequence 1 are skipping steps and short cut the individual testing of 17 differences. In the first step, the largest critical value, $R'_{25} = 12.80$ is subtracted from the largest adjusted total 39.5 (for K) giving 26.70. From this it is concluded that all treatments with adjusted totals below 26.70, namely D, P, $\cdots$, S, are significantly lower than K. They can thus be dropped from the set leaving the subset $UA \cdots K$ for testing in the second step.

The truth of this conclusion is readily seen as follows:—Consider any one of the differences concluded significant, say $K - M$ for example. We have $M(.447) < K(.447) - R'_{25}$ implying $(K - M)(.447) > R'_{25}$, thus $(K - M)(.471) > R'_{13}$, since $R'_{25} > R'_{13}$, and hence $(K - M)' > R'_{13}$. Similarly each of the adjusted differences concerned exceeds its corresponding critical value and all are thus significant.

In the second step testing the 8-treatment subset $UA \cdots K$ in a similar way, the largest critical value R'_8 involved is subtracted from the top adjusted total giving 27.4. If there were any further adjusted totals below 27.4 the treatments concerned could be dropped off and another similar step would be applied. In these data no further treatments can be dropped and the skipping method terminates at this second step. The remainder of the sequence is finished by steps of the type already described in the first example, and for which the arithmetical details are given in Table 4. Thus in step 3, $a_{++}(K - U) > R'_8$, and in step 4, $a_{++}(K - A) \not> R'_7$. This terminates the sequence since $K - A$ has the larger adjustment factor a_{++} and no other difference within the subset $AV \cdots K$ can exceed R'_7. This result may be usefully recorded as before by putting the subset in parentheses as shown in the result column of section (d) at the end of sequence 1.

TABLE 4

ARITHMETICAL DETAILS FOR CALCULATING ADJUSTED DIFFERENCES IN TABLE 3, SECTION (d)

$$(K - U)' = a_{++}(K - U) = .471(88.4 - 61.4) = .471(27.0) = 12.72.$$

Similarly

$$(K - A)' = .471(21.8) = 10.27, \quad (B - E)' = .471(26.4) = 12.43,$$
$$(O - J)' = .471(28.5) = 13.42, \quad (Q - S)' = .471(24.5) = 11.54.$$

Sequence 2 is very similar to sequence 1. The skipping procedure starts by subtracting R'_{14} from the second highest adjusted total 34.6 for B and terminates again in the second step. The remainder of the sequence terminates at the third step with $EL \cdots B$ being found homogeneous.

In sequence 3 an additional treatment is dropped in the second step and the skipping procedure extends to the third step. Continuing the remainder of the sequence, $a_{--}(O - J)$ is found to exceed R'_{15} in the fourth step, $a_{--}(O - F)$ is known not to exceed R'_{17} from the preceding skipping steps so $FGEL \cdots O$ is recorded at the fifth for further internal testing. The largest $(O - ?)$ difference with an a_{--} adjustment factor is $O - E$. However $a_{++}(O - E)$ cannot exceed R'_{17} since $a_{++}(B - E) > R'_{15}$ in sequence 2 hence $FG \cdots O$ is homogeneous and the sequence terminates. Sequence 4 has only two skipping steps and terminates in a similar way. Sequences 5, 6 and 7 each terminate rapidly at the first step when the subsets concerned are found to fall entirely within $(CH \cdots N)$ of sequence 4. In sequence 8, the difference between the adjusted total for I, that is, 17.0 and the critical level 14.67 is so great as to leave no doubt of the final result $(IC \cdots U)$ at the end of the first step.

As soon as all treatments but the lowest have been included in homogeneous subsets, as is the case in the example at the end of sequence 8, the test can be completed in one reverse type of sequence working from the bottom total. The reasons for this will be evident from the steps of the last sequence in section (d). The largest possible homogeneous subset in which the bottom treatment S could be included at this stage is $SI \cdots D$ and contains 17 treatments. The first step is to obtain $S + R'_{17} = 9.6 + 12.69 = 22.29$. From this it follows that all treatments with adjusted totals above 22.29, that is E, L, T, $\cdots$ are significantly larger than S. This leaves the 8-treatment subset $SI \cdots G$ for testing in the next step. In the second step $S + R'_8 = 21.70$, no additional adjusted totals exceed this and the skipping procedure terminates. Since the range $G - S$ of the subset has the adjustment factor a_{+-} we already know from the step 2 that $a_{+-}(G - S) > R'_8$ hence $SICHOJFG$ is recorded for internal testing. The largest $(? - S)$ difference with the a_{--} adjustment factor is $(Q - S)$ and this is therefore tested in step 3. Since $a_{--}(Q - S) > R'_2$ the subset is homogeneous, and is recorded ($SICHOJFG$). This terminates the test.

Section (e) of Table 3 shows the complete summary of the test results. In this example the ordering of treatments does not vary from one homogeneous subset to another. In such a case the method of representing the results by underscoring a single set of treatment letters as is done in [3] may be used if preferred.

5. NOTES ON THE PROPERTIES OF THE PROPOSED TEST

Two-mean significance levels: A two-mean significance level in a test of n means may be defined [3] as the maximum probability of finding a significant difference between any two means m_i and m_j given that $\mu_i = \mu_j$ where $\mu_i = E(m_i)$ and $\mu_j = E(m_j)$. This may be written as the $\max \cdot P[D_{ij} \mid \mu_i = \mu_j]$ where D_{ij} denotes the decision that m_i and m_j are significantly different.

In any α-level test of the proposed type we have $\max \cdot P[D_{ij} \mid \mu_i = \mu_j] = P[a_{ij} \mid m_i - m_j \mid > s \cdot z_{2,n_2,\alpha} \mid \mu_i = \mu_j]$. Since $z_{2,n_2,\alpha} = \sqrt{2}\, t_{n_2,\alpha}$ where $t_{n_2,\alpha}$ is the α-level (two-sided) significant value of t_{n_2}, a t statistic with n_2 degrees of freedom, and since the variance of $m_i - m_j$ is $2\sigma^2/a_{ij}^2$, this readily reduces to

$$P[\mid t_{n_2} \mid > t_{n_2,\alpha}] = \alpha.$$

Hence the two-mean significance levels in an α-level test of the proposed type are exactly α as desired.

Higher order significance levels and power: In considering these further aspects of the proposed test it is helpful to study the decision regions for a 5 per cent level test [3] of three unequally replicated means m_1, m_2 and m_3 with $r_1 = 2$, $r_2 = 3$ and $r_3 = 4$ replications and in which $n_2 = \infty$ and $s^2 = \sigma^2 = 1$. If these regions are plotted in a plane with coordinates

$$x_1 = (m_1 - m_2)\sqrt{2r_1r_2/(r_1 + r_2)}$$

and

$$x_2 = (r_1m_1 + r_2m_2) - (r_1 + r_2)m_3\sqrt{2r_3/(r_1 + r_2)(r_1 + r_2 + r_3)}$$

as is done in Figure 1, they are directly comparable to those of the corresponding 5 per cent level test of three equally replicated means ($r_1 = r_2 = r_3$) shown in Figure 3 of [3]. The distribution function for the points (x_1, x_2) is the same in both cases, namely a bivariate normal with variances 2 and 2 and with covariance zero.

Because the strip regions (1, 2), (1, 3) and (2, 3) have the same minimum widths in each case ($2z_{2,\infty} = 2 \times 2.77$) it follows, as has already been proved, that the two-mean significance levels are 5 per cent for the test in Figure 1 as well as for the test with equal replications.

The only differences between the regions in the two cases is that in Figure 1 the angles between the strip regions (1, 2), (1, 3), (2, 3) and (1, 2) are 50°46′, 67°48′ and 61°26′ instead of all being 60° as in the other figure. (The cosine of the angle between any two strips (h, i) and (h, j) is given by

$$\sqrt{r_ir_j/[(r_h + r_i)(r_h + r_j)]}).$$

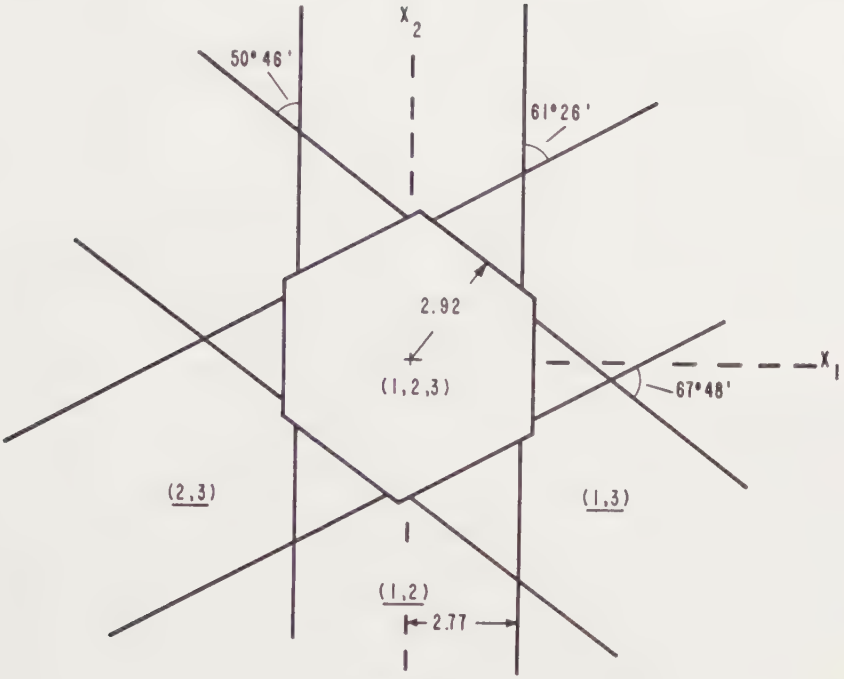


FIGURE 1

5% level test, $r_1 = 2$, $r_2 = 3$, $r_3 = 4$, $n_2 = \infty$, $\sigma^2 = 1$.

The sides of the hexagonal regions $(\underline{1, 2, 3})$ are parallel with the corresponding strips. Since these hexagons have the same inscribed circle of radius $z_{3,\infty} = 2.92$ and differ only in having a little asymmetry in Figure 1 the three-mean protection level (the probability $P[(x_1, x_2) \in (\underline{1, 2, 3}) | E(x_1) = E(x_2) = 0]$) for the Figure 1 test is a close approximation to the desired level .9025 obtained in the other test. Furthermore, it seems safe to assume that any deviation due to the asymmetry would be positive.

In terms of three-mean significance levels, the level of the Figure 1 test may thus be said to be close to .0975 (= 1 − .9025) as desired and that any deviation from this appears to be on the negative or conservative side.

The close similarity of the regions of Figure 1 with those of Figure 3 [3] also indicates that the power functions of the Figure 1 test closely approximate the desirable ones of the other procedure.

If the means are correlated as well as being heteroscedastic the geometrical picture is virtually unchanged. If we let $[c_{ij}]_{3\times 3}$ represent

the dispersion matrix in the case of three means the cosine of the angle between any two strips (h, i) and (h, j) in a set of regions otherwise similar to those of Figure 1 would be

$$(c_{hh} - c_{hi} - c_{hj} + c_{ij})/(\sqrt{(c_{hh} - 2c_{hi} + c_{ii})(c_{hh} - 2c_{hj} + c_{jj})}$$

This may be expressed in terms of the k_{ij} factors ($k_{ij}\sigma^2 = c_{ii} - 2c_{ij} + c_{jj}$) defined in section 2, as

$$(k_{hi} + k_{hj} - k_{ij})/2\sqrt{k_{hi}k_{hj}}$$

and the degree of asymmetry involved depends on these.

Similar considerations lead to the conclusion that the higher order levels of the proposed complete test and its power functions are reasonably close to the desired levels and functions existing in corresponding tests of uncorrelated and homoscedastic means and that the deviations involved appear to be on the conservative side.

REFERENCES

[1] Bleicher, Edwin [1954]. *The Extension of the Multiple Comparisons Test to Lattice Designs*, Virginia Polytechnic Institute M.S. thesis, 55 pp.

[2] Cochran, W. G. and Cox, G. M. [1956]. *Experimental Designs*, 2nd Ed., John Wiley and Sons, Inc., New York.

[3] Duncan, David B. [1955]. Multiple Range and Multiple *F* Tests, *Biometrics*, *11*, 1–42.

[4] Federer, Walter T. [1955]. *Experimental Design*, Macmillan Company, New York.

[5] Keuls, M. [1952]. The Use of the 'Studentized Range' in Connection with an Analysis of Variance, *Euphytica*, *1*, 112–122.

[6] Kramer, Clyde Young [1956]. Extension of Multiple Range Tests to Group Means with Unequal Numbers of Replications, *Biometrics*, *12*, 307–310.

[7] Kramer, Clyde Young [1957]. Extension of Multiple Range Tests to Group Correlated Adjusted Means, *Biometrics*, *13*, 13–18.

[8] Newman, D. [1939]. The Distribution of the Range in Samples from a Normal Population, Expressed in Terms of an Independent Estimate of Standard Deviation, *Biometrika*, *31*, 20–30.

[9] Sanders, Paul G. [1953]. *The Extension of the Multiple Comparisons Test to Rectangular Lattices*, Virginia Polytechnic Institute M.S. thesis, 31 pp.

[10] Tukey, J. W. [1953]. The Problem of Multiple Comparisons, unpublished dittoed notes, Princeton University, 396 pp.

APPROPRIATE SCORES FOR REACTION CATEGORIES DEPENDENT ON TWO VARIABLES*

JOHANNES IPSEN, JR.

Associate Professor of Public Health, Harvard School of Public Health, Superintendent, Massachusetts Institute of Laboratories, U.S.A.

1. In an earlier paper (Ipsen [2]) the appropriate reaction scores for bio-assays were given where the reactions to various doses of a biologic agent were observed in biological, mutually exclusive categories. Death times and survivor symptoms were examples of such categories. Reactions to anti-emetic drugs can be similarly treated (Ciminera et al. [1]).

For such biometric purposes the appropriate scores are defined as the scores whose linear regression on the independent variable (dose) is a maximum part of the total variation of reaction scores. They are computed as the mean dose for each category or any linear transformation thereof.

Thus, if n_i subjects that fall in the i-th category of reaction are exposed to various doses, the sum of which is $S(x_i)$, the appropriate score (c_i) for that category would be

$$c_i = b\,\frac{S(x_i)}{n_i} + a \tag{1.1}$$

where b is different from zero and a can be any rational number.

The choice of a and b does not affect the information.

2. If the subjects are exposed to two agents with doses x and z respectively, the problem of appropriate scores will consist in assigning a score system that will maximize the multiple regression contribution to the total variance.

The following notation will be used:

N = Total number of subjects

*This work was conducted under the sponsorship of the Commission on Immunization Armed Forces Epidemiological Board, and was supported in part under contract with the Office of the Surgeon General, Department of the Army.

n_i = Number of subjects in the i-th category

$S(x_i)$, $S(z_i)$ = Sum of respective doses in the i-th category.

$S(x)$, $S(z)$ = The total sum of respective doses

$$[x^2] = S(x^2) - (S(x))^2/N$$

$$[z^2] = S(z^2) - (S(z))^2/N$$

$$[xz] = S(xz) - S(x)S(z)/N$$

$$[x_i^2] = \sum_i S^2(x_i)/n_i - (S(x))^2/N$$

$$[z_i^2] = \sum_i S^2(z_i)/n_i - (S(z))^2/N$$

$$[x_i z_i] = \sum_i S(x_i)S(z_i)/n_i - S(x)S(z)/N$$

The appropriate score system (y_i) for the multiple regression is linearly related to each of the appropriate systems for the single regressions.

$$y_i = (S(x_i) + \beta S(z_i))/n_i \tag{2.1}$$

or linear transformations thereof.

Consequently, we have

$$S(y - \bar{y})x = [xy] = [x_i^2] + \beta[x_i z_i] \tag{2.2}$$

$$S(y - \bar{y})z = [zy] = [x_i z_i] + \beta[z_i^2] \tag{2.3}$$

$$S(y - \bar{y})^2 = [y^2] = [x_i^2] + 2\beta[x_i z_i] + \beta^2[z_i^2] \tag{2.4}$$

The multiple regression sum of squares is

$$R = \frac{[xy]^2[z^2] + [zy]^2[x^2] - 2[xy][zy][xz]}{[x^2][z^2] - [xz]^2} \tag{2.5}$$

$\theta = R/[y^2]$ is then the fraction which should be maximized by a suitable choice of the coefficient β. (2.6)

Inserting (2.2), (2.3) and (2.4) in the expression for θ, differentiating θ with respect to β and equating the result to zero, we have

$$\begin{aligned}\beta^2([z_i^2][xz] - [x_i z_i][z^2]) + \beta([z_i^2][x^2] - [x_i^2][z^2]) \\ + [x_i z_i][x^2] - [x_i^2][xz] = 0\end{aligned} \tag{2.7}$$

The two solutions for β inserted in (2.1) will give two score systems that represent a maximum and minimum regression *variance contribution*, respectively. It is easy to ascertain which of the two values of β will yield maximum information.

3. *Example. Untoward reaction to combined diphtheria and tetanus immunization.* The modern trend in the practice of immunization is to combine several antigens in one preparation. The advantage of fewer injections is, however, sometimes offset by the higher frequency of untoward reactions such as local swelling, fever, and other discomforts. This is particularly true in adults where it is often found that the degree of untoward reaction is related to the size of antibody response. Diphtheria and tetanus toxoids have been combined for a long time, and it has usually been assumed that untoward reactions to such combinations are solely due to previous exposure to diphtheria toxin or other less defined antigens in the prophylactic. The more the individual has been exposed to diphtheria antigen either through natural infections or through immunization, the higher is the ensuing antibody production as well as the probability of discomfort. Tetanus immunity is not acquired by natural infection and it has long been observed that a single injection of tetanus toxoid rarely induces reaction. However, in the last two decades more and more people have received several injections of tetanus toxoid, and it is to be expected that greater sensitivity to this antigen will occur.

42 young adults were each give one injection of a combination of diphtheria and tetanus toxoid that contained 1 Lf unit of diphtheria and 5 Lf units of tetanus toxoid (Ipsen [3]). Reactions were recorded in the days following the injections and a blood sample was drawn four weeks after the injection for measurement of antibody concentration in the serum. It was estimated that about one-half of these individuals had previous exposure to either diphtheria or tetanus antigen or both. Table 1 presents the serum antibody titers for the two agents, arranged in pairs, and in four categories of observed reaction. The antibody titers are given in logarithms of antitoxin units per 100 ml. The titrations for tetanus were not carried below .1 unit per ml. (or 10 units per 100 ml.). Specimens with less than this amount are recorded as 1.0. For diphtheria the lowest observation was 0.01 units per ml. (or 1 unit per 100 ml.) Specimens with this amount or less are recorded as 0.0.

The biometric problem consists in assigning a score system to the reaction that will give maximum information on the relation of reaction to both antibody titers simultaneously. The biological assumption is that such reaction is positively correlated with the amount of antibody response in both cases. The bottom of Table 1 presents the statistics used for the computation of the co-efficient β that determines the relative importance of the two antigens. Inserting statistics in equation (2.7), we obtain the following expression:

$$65.633\beta^2 + 263.309\beta - 6.576 = 0 \qquad (3.1)$$

of which the two solutions are

$$\beta = \begin{cases} +0.025 \\ -4.037 \end{cases}$$

TABLE 1

Diphtheria (z) and tetanus (x) serum antitoxin in 42 adults, 4 weeks after one injection of combined diphtheria and tetanus toxoids, by untoward reaction category. (Titers are in log antitoxin units per 100 ml. serum.)

(*A*) None		(*B*) Local redness or soreness		(*C*) Malaise, fever etc. "systemic"		(*D*) Local and systemic	
Diph.	Tet.	Diph.	Tet.	Diph.	Tet.	Diph.	Tet.
0.0	1.0	0.0	1.0	0.8	1.0	1.3	2.8
0.0	1.0	0.5	1.0	0.0	2.8	1.1	3.4
0.0	1.0	2.0	1.0	1.3	2.8	2.5	3.4
0.0	1.0	0.0	1.3	0.0	3.1		
0.2	1.0	0.8	1.3	0.3	3.4		
0.2	1.0	1.1	1.9	0.8	3.4		
0.5	1.0	1.6	1.9	1.6	3.4		
0.8	1.0	0.0	2.2				
0.8	1.0	0.0	2.2				
1.1	1.0	0.0	2.5				
1.9	1.0	2.8	2.5				
2.0	1.0	1.3	2.8				
0.0	1.3	2.1	3.4				
1.1	1.6						
0.0	2.2						
0.0	2.8						
1.2	2.8						
1.9	2.8						
1.1	3.4						

	A	*B*	*C*	*D*	Totals
n_i	19	13	7	3	42
$S(x_i)$	28.9	25.0	19.9	9.6	83.4
$S(z_i)$	12.8	12.2	4.8	4.9	34.7

$S(x^2) = 203.28$ $\quad$ $S(z^2) = 55.27$

$[x^2] = 37.6714$ $\quad$ $[x_i^2] = 13.7196$

$[z^2] = 26.6012$ $\quad$ $[z_i^2] = 2.6983$

$[xz] = 8.7257$ $\quad$ $[x_i z_i] = 3.3524$

The two score systems are now computed by inserting β into equation (2.1), the values for which are arranged in Table 2. However, it is more convenient to convert this system to linear transformations

$$y_i' = (y_i - y_A)/(y_D - y_A) \tag{3.2}$$

which have score 0 for the lowest category (A) and score 1 for the highest category (D) as shown in Table 2. The choice between these score systems depends on which system will give the highest information measured by θ. Using the equations (2.2) to (2.6) we obtain

$$\text{for } \beta = \quad 0.025; \qquad \theta = 0.3643$$

$$\text{for } \beta = -4.037; \qquad \theta = 0.0756$$

Obviously, then, the positive value of β is the desired co-efficient since it gives the highest information. Also the score system is more "logical" since it follows the biological ranks of the categories.

TABLE 2

Score Systems with Maximum and Minimum Information

Category	$\beta = 0.025$		$\beta = -4.037$	
	Y_i	Y_i'	Y_i	Y_i'
A	1.538	0.00	−1.199	0.00
B	1.947	0.24	−1.866	0.30
C	2.860	0.78	+0.075	−0.58
D	3.241	1.00	−3.394	1.00

4. *Biological Conclusions*: Using the appropriate score system, we can now express the expected degree of reaction on the basis of a known antibody response to diphtheria and tetanus respectively, by means of the multiple regression equation

$$y' = -0.1545 + 0.2146x + 0.0050z \tag{4.1}$$

It is obvious that the influence of tetanus antitoxin is much greater on the response than that of diphtheria antitoxin. As a matter of fact, the information obtained by tetanus antitoxin alone amounts to 0.3641 while the maximum information obtained by combining the two independent variants amounted to 0.3643. The difference of less than 1 per thousand is then the estimated additional information that would come from including the diphtheria antibody response in the expectancy equation.

This analysis has prompted our Biological Laboratories to conduct further investigations into the importance of the tetanus component of combined vaccines in respect to untoward reactions.

The author wishes to acknowledge the assistance of Harry E. Bowen, Ph.D., in arranging the immunization trial, and to Mrs. Hanna Sylwestrowicz, B.A., for performing the antibody titrations.

REFERENCES

[1] Ciminera, J. L., Stone, C. A., and Ipsen, J. [1957]. The Statistical Evaluation of a Biological Assay for Anti-Emetic Drugs. *Sc. Ed. J. Am. Pharm. Assoc. 46*: 197–200.

[2] Ipsen, J. [1955]. Appropriate Scores in Bio-Assays Using Death-Times and Survivor Symptoms. *Biometrics 2*: 465–480.

[3] Ipsen, J. [1954]. Immunization of Adults Against Diphtheria and Tetanus. *New Eng. J. Med. 251*: 459–466.

POLYMORPHISM IN SOME AUSTRALIAN LOCUSTS AND GRASSHOPPERS

R. E. Blackith
Imperial College Field Station, Sunninghill, Berks., England

INTRODUCTION

The measurement of changes of form in closely related organisms has often been attempted by compounding pairs of characters into ratios, or by examining those transformations which secure the superposition of outlines of the organisms traced on some system of coordinates (Le Gros Clarke and Medawar [1945]). Attempts to make the coordinate transformation method quantitative, and the ratio method more general, have not been conspicuously successful. The limitations of ratios and their generalisation are illuminated by comparing them with discriminant functions in which characters are added or subtracted. A change of scale, to a logarithmic coordinate network, will transform the ratio A/B to the discriminant $\log A - \log B$. Here there are two characters, with equal and opposite weights, but the discriminant function could evidently be constructed of some such expression as $\hat{X} = \hat{\beta}_1 \log A - \hat{\beta}_2 \log B + \hat{\beta}_3 \log C \cdots$ which is restored, on relaxing the scale distortion, to $\hat{Y} = (A^{\hat{\beta}_1} \cdot C^{\hat{\beta}_3}/B^{\hat{\beta}_2})$, which also serves to generalise the classical equation of allometric growth $\hat{Z} = k \cdot A^{\beta}$ (Teissier [1937]). Any number of characters may be included in this way, the weights being calculated to give the greatest possible discrimination between any pair of groups of organisms under comparison.

The use of linear discriminant functions of the form $X = \beta_1 A + \beta_2 B + \beta_3 C + \cdots$ for the assessment of differences of form and size in quantitative anthropology, (Mahalanobis, Majumdar and Rao [1949]) and in the analysis of the changes of shape of silkworm cocoons, (Fraisse and Arnoux [1954]) to name only two examples with animals, has had sufficient success to suggest that a quantitative appraisal of changes of size and shape has in this way become available almost unnoticed by those interested in the classical problems of growth and form. These discriminant functions are vectors which define the direction in which two groups of organisms differ, so that where more than one pair of

groups is under comparison, the contrasts between them may differ qualitatively as well as quantitatively. A change in the relative weights of the characters which make up the function alters the direction of the contrast between the groups. With each discriminant function there is associated a corresponding separation of the two groups. This disjunction is defined by the expression

$$D^2 = \beta_1 d_1 + \beta_2 d_2 + \beta_3 d_3 + \cdots + \beta_i d_i + \cdots$$

where D is the generalised distance between the groups. This distance separates the mean position of each group in a space of as many dimensions as there are characters in the discriminant function. As for the rest of the expression, the β-coefficients are the weights of the characters in the discriminant function. These coefficients are found by serially multiplying the differences (d_i) between the mean values of each character in the two groups by the rows or columns of the inverse of the dispersion matrix. The original dispersion matrix (**A**) consists of the variances and covariances of the characters, arranged in matrix form:—

$$\mathbf{A} = \begin{matrix} a_{11} & a_{12} & a_{13} \cdots \\ a_{21} & a_{22} & a_{23} \cdots \\ a_{31} & a_{32} & a_{33} \cdots \\ \cdots\cdots\cdots\cdots \end{matrix}$$

in which such entries as a_{11} along the leading diagonal represent the variances of the several characters, and the entries such as a_{12} represent the covariances, for example of the first and second characters. The determinant of this dispersion matrix has been described by Kendall [1946] as playing the same part in multivariate analysis as does the variance in the ordinary, univariate, case. The inverse of this matrix, $\mathbf{A}^{-1}$, may be written out with entries c_{ij} in place of the a_{ij} of the dispersion matrix **A**. These entries in $\mathbf{A}^{-1}$ are expeditiously found by the method described by Rao [1952] but there is a variety of computational arrangements advocated by different authors, for what is essentially the process of disentangling the correlations between the several characters. This process of finding the β-coefficients is also used in multiple regression studies, in which the adjusted sums of cross-products between the dependent and independent variates replace the d's of the present case. The further multiplication of the β-coefficients by the d_i's gives the square of the generalised distance.

Thus for each discriminant function there is an associated distance, so that when the separation of two groups of organisms is made as large as possible, for a given set of characters in the function, these

two groups will be separated in a definite direction in the multi-dimensional space. We may expect that, when biologically similar contrasts are made between groups of organisms (as for instance, between the sexes of several related species) these contrasts will be similarly oriented in the common hyperspace, even though the generalised distances between the sexes, i.e. the degree of sexual dimorphism, may differ from species to species. These vector properties of discriminant functions have been so little used since Fisher [1938] described their application as of urgent importance twenty years ago, that Goodall [1954] is almost the only author to make explicit use of them, in this instance in phytosociological investigations. The arrangement of groups of organisms, separated by the generalised distances between each pair of groups, in an appropriate hyperspace has been called discriminatory topology, and the descriptions of the method accompany the anthropometric investigations of Mahalanobis, Majumdar and Rao [1949], of Rao [1952] and of Mukherjee, Rao, and Trevor [1955], all of whom used generalised distances mainly as indicators of group separation, as did Hughes and Lindley [1955], without paying much attention to the direction of separation.

There is, nevertheless, a closely related method described by Rao [1952], by which the several groups may be depicted on a chart, so that the underlying relationships between their forms can be exhibited. Representing, as before, the pooled dispersion matrix as **A**, we introduce the matrix **B** which represents the dispersion of the different groups in the hyperspace just as **A** represents that of the individual organisms about the mean for their group, the determinantal equation:

$$| \mathbf{A} - \lambda \mathbf{B} | = 0$$

may be solved to give as many solutions for the latent roots as there are characters measured. Each of these roots is associated with a vector, which generates an axis along which the position of the group may be plotted. Each of the vectors or canonical variates associated with the latent roots is orthogonal, so that, provided that only the first three roots account for important fractions of the disparity between the groups, a solid model may be made of their mutual relationship, and, for two important dimensions, a planar representation is possible. This method should not be confused with the evaluation of the latent roots of the characteristic equation of the dispersion matrix,

$$| \mathbf{A} - \lambda \mathbf{I} | = 0$$

where **I** is the corresponding unit matrix with each entry in the leading diagonal unity and the remainder zero. The roots are then the latent

roots of the dispersion matrix (the eigenfunctions of statistical mechanics) and have associated with them the principal axes of the ellipsoid of individual points scattered about the general mean. The extraction of these principal components has close affinities with factor analysis (Burt and Banks [1947]; Teissier [1948]).

AN ILLUSTRATIVE EXAMPLE

To illustrate the practical application of the approach outlined above the discriminatory topology of some Australian locusts and grasshoppers is presented.

Locusts and grasshoppers are noted for their marked plasticity of form, colouration and behaviour. The density at which locusts are reared modifies their form, and that of their progeny (Albrecht [1955]).

The limiting structural types then represent the phases of locusts, (Uvarov [1921]) which may be distinguished from other and density independent phase phenomena as kentromorphic phases (Key and Day [1954]). The insects which develop from persistently crowded populations constitute the *gregaria* phase, those from dispersed populations fall into the phase *solitaria*. The relations between the density and possible swarm formation are complicated, but the possibility of the prediction of swarming is so important that much effort has gone to discover ways of measuring locusts so that the cumulative influence of population density may be assessed. Usually, pairs of measurements are compounded in ratios, and this empirical usage has tended to obscure the intention of allometric studies to distinguish the underlying modes of growth of different parts of an organism. The object of this paper is to present the relationships which can underlie, not so much the measurements made on a homogeneous group of locusts or grasshoppers, as those made on different groups. The mutual relations of the form of different phases, species and sexes of these insects have important biological implications.

The technique used here is to compute the generalised distances between every pair of groups. These distances may then be drawn to scale on a chart or as a three dimensional model, with the same limitations as noted above for canonical variates.

In such generalised distance charts the vector properties of discriminant functions (Fisher [1938]) reveal underlying dimensions along which changes of external form can be assessed. The generalised distances between the two sexes of each species may be parallel to one another, but not to those generalised distances which link insects of the same sex but of different phases, unless phase differences are no

more than that exaggeration of normal growth reflected in sexual dimorphism. The method may be extended to include differences of form between subspecies or higher taxa, provided that the measurements are homologous, and to include groups of insects with different ecological backgrounds. Such insects may reflect in their shapes the consequences of living in different habitats, as the 'variation écologique' of Pasquier [1938].

Just as studies of allometric growth have pioneered the investigation of underlying relationships among measurements made on a homogeneous group of organism, (Huxley [1932], Teissier [1937]) so comparisons between the shapes of different groups of organisms become possible. Notwithstanding the qualitative methods of D'Arcy Thompson and their ramifications (Le Gros Clark and Medawar [1945]) the development of quantitative techniques for inter-group comparisons of form has followed different lines of thought (Cousin [1948]: Anderson [1953]). Such methods often stem from the concepts of factor analysis and there is evidence that factor analysis and principal component analysis, applied to the variation of form in arthropods, yield essentially similar results (Teissier [1955a]). The extension of techniques using generalised distances to include differences of taxonomic rank is straightforward, leading to a three-dimensional structure if the differences between the sexes, which are primarily a matter of size, and those between the phases or the species or subspecies are each qualitatively distinct reflections of underlying modes of growth. Such virtually orthogonal relationships between the representation of size, phase, and specific variation have been found by Albrecht and Blackith [1957] among some African species of locust.

Such relations are found by inspection of the generalised distance charts after all the groups of locusts have been located. The emphasis in this investigation concerns the vector, rather than the scalar, properties of the individual comparisons between the groups of locusts and grasshoppers. Not only are generalised distances appropriate for this purpose, but by concentrating attention on the direction rather than the magnitude of the separation afforded one evades the problem of whether the material studied exhibits the full range of morphological plasticity of which it is capable.

EXPERIMENTAL MATERIAL

This paper reports a new analysis of the measurements of some Australian locusts and grasshoppers given by Key (1954), who has used methods customary in the study of locust morphology, namely the

comparison of individual characters and their ratios. The characters employed here are listed in Table 1.

Two genera are represented, *Chortoicetes* Brunn. and *Austroicetes* Uv. of which the latter has several species, some rare, one of which is further divided into subspecies.

A. arida Key (10♂♂, 7♀♀)
A. vulgaris vulgaris Sjöst. (10♂♂, 10♀♀)
A. vulgaris corallipes Sjöst. (9♂♂, 10♀♀)
A. nullarborensis Key (30♂♂, 36♀♀)
A. cruciata Sauss. (29♂♂, 29♀♀)
A. tricolor Sjöst.. (5♂♂, 10♀♀)
A. pusilla Walk. (10♂♂, 10♀♀)
A. tenuicornis Key (1♂, 2♀♀)
A. frater Brancs. (10♂♂, 10♀♀)

A. nullarborensis and *A. cruciata* exhibit the phase polymorphism commonly found in locusts, that is to say their form and some features of their colouration are density dependent. They also exhibit non-phase colour polymorphism and sexual dimorphism in addition to phase polymorphism, but as more suitable material is available elsewhere for assessing the morphometric consequences of colour polymorphism this topic will not be pressed here.

TABLE 1

SEPARATION (D^2-VALUES) OF THE PHASES BY DISCRIMINANT FUNCTIONS COMPOUNDED OF 1–6 CHARACTERS

Species:—	*A. cruciata* (Eastern race)		*A. nullarborensis.*	
Character	Males	Females	Males	Females
Elytron length	8.25	0.07	3.72	0.50
Posterior femoral length	8.27	0.08	3.73	0.73
Headwidth	11.83	0.25	4.25	1.81
Pronotal length	12.98	3.43	4.80	1.85
Pronotal width	12.98	3.55	5.40	1.92
Pronotal height	12.99	3.58	5.50	1.92

The single species of *Chortoicetes*, *C. terminifera* (Walk.) has two races, morphologically distinguishable, which inhabit respectively South-Western Australia and most of the remainder of the continent. These forms are called the 'southwestern' and 'eastern' races, and both exhibit phase polymorphism. Complete sets of 6 measurements of

64♂♂ and 68♀♀ *Chortoicetes* were available for analysis and the available numbers of *Austroicetes* are given in parentheses after the specific name, though Key had many other incomplete sets on which his conclusions are based. For several rare species the numbers are small; many will prefer to discount such tentative results, others may prefer to see how even rare material may be used to give tentative results, which process is, after all, no more than usual taxonomic practice. The main interest of the analysis centres on the relatively abundant *C. terminifera*, *A. nullarborensis* and *A. cruciata*.

METHODS OF ANALYSIS

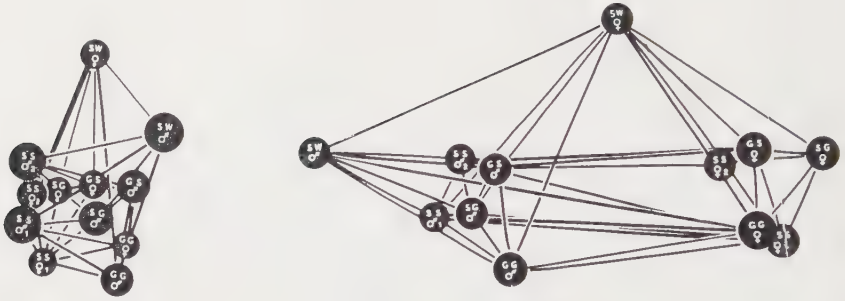
The method used here was to prepare dispersion matrices for each species for which 10 or more of each sex-species-race-phase category has been measured. These matrices were then pooled, and the inverse of this pooled matrix was linked to the differences between the mean values of the pairs of categories as described in the introduction. From the generalised distances so computed the charts of Figs. 1 and 2 were prepared. Fig. 1 is a projection, Fig. 2 a photograph, of the models made by cutting glass rods proportional in length to the generalised distances, and fitting these rods into rubber balls representing the groups of insects.

The fact that the models form a three-dimensional framework, whereas the hyperspace is six-dimensional, may be taken to indicate that not more than three underlying dimensions of variation are important. The discrepancies in the positions of the balls located by more than one set of generalised distances were small enough to be considered as sampling variation, and accommodated by adjusting the position within the balls at which each rod terminated.

RESULTS

We have to identify the three important dimensions of variation within the charts. The general 'sex' dimension is well established and even the rare species for which few individuals are available conform to it. This reflection of sexual dimorphism is essentially a distinction of size. The segregation of the sexes along this dimension is well brought out in the figures, the females being uniformly larger than the males in the species studied.

In these charts, there is a marked three-dimensionality, and we have to identify, if possible, the remaining two sources of variation other than size. Density-dependent changes of form, as between the phases of locusts, are known to constitute one such dimension of varia-



(a) (b)

FIGURES 1a AND 1b.

GENERALISED DISTANCE CHART FOR *Chortoicetes terminifera*, BASED ON SIX CHARACTERS.

a) END VIEW OF CHART, LOOKING ALONG 'SIZE' AXIS.
b) SIDE VIEW OF CHART LOOKING ALONG 'PHASE' AXIS.

SOUTH WESTERN RACE.

SW *C. terminifera*

EASTERN RACE.

SS_1 *C. terminifera* from cooler, moister areas where swarms never occur.
SS_2 Solitary insects of Eastern race, probably of solitary parentage, but from warmer, drier area. (SS_1 and SS_2 are of ph. *solitaria*).
SG Solitary insects, probably from gregarious parents.
GS Gregarious insects, from solitary parents.
GG Gregarious insects, from gregarious parents. (ph. *gregaria*).

tion (Albrecht [1955]). Within the genus *Austroicetes* the planes defined by the generalised distances between the phases and sexes of *A. nullarborensis* and *A. cruciata* (Eastern race) are nearly parallel. Moreover, the generalised distances which link corresponding groups of the two species are roughly at right angles to those planes linking the phases and the sexes, showing that the interspecific difference of form is not an exaggeration of sexual dimorphism or phase variation to be found in the common ancestor.

The chart for the genus *Chortoicetes* (Fig. 1a & b) reveals that the 'phase' dimension of the chart is delimited by groups of locusts which are, respectively, reared in comparative isolation and bred from dispersed parents, or reared in crowds and bred from crowded parents. There

are also two intermediate categories; locusts from crowded parents but themselves reared in comparative isolation, and locusts whose parents were dispersed but which, themselves, have been reared crowded. These intermediate forms lie between the two extremes (Fig. 1a). The inference is that the full morphometric expression of phase status is not attained unless the parental density accords with that of progeny, and both are at extremes of the normal biological range. On the evidence of this investigation, the influence of parental isolation, working in the opposite sense to that of the crowding of the immature stages, reduces the phase status to a position about half-way between the extremes as represented in Fig. 1a. Quantitatively, this conclusion is tentative until the cumulative densities of parent and progeny populations can be established in the field with more accuracy than has been possible so far. This finding agrees with those of Albrecht [1953] (and private communication) who found that the parental density of Red and Migratory Locusts modifies the form of progeny so that unless both the parents and the progeny are crowded in immature life the latter will not attain full *gregaria* phase status.

The charts help to explain anomalies arising from the use of ratios of characters. For instance, Key [1954] points out that the South-Western race of *Chortoicetes terminifera*, assessed in terms of the ratio elytron length/posterior femoral length, appears to be 'super-*solitaria*', in the sense that this ratio takes values for that race which are normally associated with highly dispersed populations of *Chortoicetes* even though the populations were in fact far from sparse. On the other hand, other ratios, of head and pronotal characters, gave contradictory results, and Key indicates that the differences between the races must be distinct from those which exist between phases. Fig. 1(a & b) shows that this South-Western race is in fact displaced from the Eastern race along a dimension qualitatively distinct from the 'phase' or 'sex' dimensions. Some interesting features emerge from attempts to identify this third dimension.

Considering only the Eastern race of *C. terminifera*, there are represented on the chart two samples of approximately the same phase status but from localities providing different habitats for the insects. The groups labelled SS_1 come from regions at once moister and cooler than the regions from which come samples SS_2. Fig. 1b supports Key's conclusion that sample SS_1 differs from the others perhaps partly through the influence of climate superimposed on differences in gregarization. We find that our third dimension encompasses alike differences of form between races and the lesser differences between insects from distinct habitats. Moreover, in Fig. 2 it is that dimension

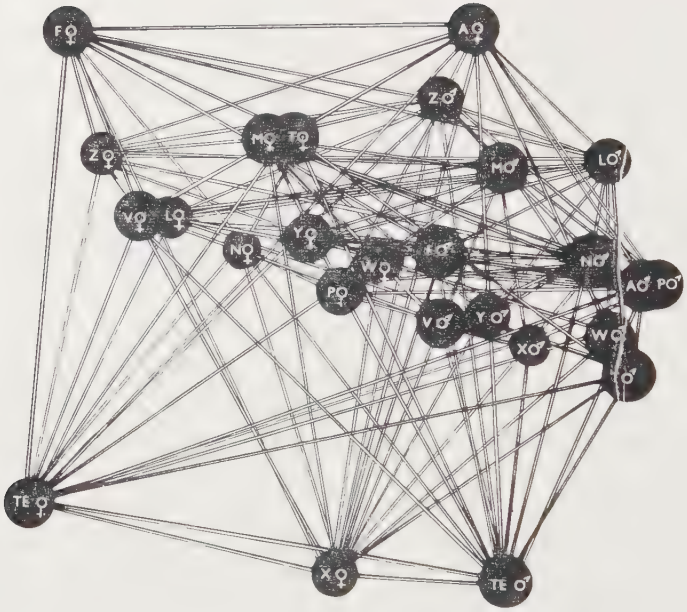


FIGURE 2

Generalised Distance Chart for *Austroicetes* spp. Based on Six Characters.

A. *Austroicetes arida*
F. *A. frater.*
L. *A. cruciata* solitarioid (Eastern race)
Z. *A. cruciata* gregarioid (Eastern race)
Y. *A. cruciata* gregarioid (Western race)
X. *A. cruciata* solitarioid (Western race)
P. *A. pusilla*
V. *A. vulgaris corallipes*
W. *A. vulgaris vulgaris*
T. *A. tricolor.*
TE. *A. tenuicornis*
M. *A. nullarborensis* gregarioid
N. *A. nullarborensis* solitarioid

normal to the 'phase' and 'sex' dimensions in which most of the specific differentiation within the genus *Austroicetes* lies just as in Fig. 1 the subspecific and ecological variation occupies this third dimension normal to those of phase and of sexual dimorphism. As the same six characters are measured for each genus, these share a common hyperspace of six dimensions, the two most important underlying dimensions being identifiable in common. This identification may be carried further by

linking the figures for *Austroicetes* and for *Chortoicetes*. In fact, the *Chortoicetes* lie underneath the *Austroicetes* with the males and females of each chart aligned. The differences between the genera are, in effect, exaggerations of those changes of form which accompany speciation and subspeciation in *Austroicetes* and ecological variation in *Chortoicetes*. This comprehensive dimension completes the identification of the three major manifestations of morphological plasticity of the species studied.

Among the *Austroicetes*, *A. arida* is rather oddly aligned, and the *solitaria* phase of *A. cruciata* (Western race) has the generalised distance between the sexes aligned with the 'species' dimension. One might suggest that these insects were hybrids, as it is known that in gryllids, at least, the inheritance of form is liable to show marked sex-linkage (Cousin [1948]), were it not that the cytological and other evidence discounts this suggestion (White and Key [1957]).

THE EFFICACY OF DIFFERENT MEASUREMENTS AS INDICATORS OF LOCUST PHASE STATUS

Much research has been directed to the discovery of characters which reveal the phase status of locusts. Anderson [1953] has emphasised the importance of a careful choice of morphometric characters in botanical work. One depends on the information provided by a discriminatory analysis to verify the wisdom of this choice.

Key's choice of characters has been assessed by computing the successive D^2-values which differentiate the phases of the Australian species he studied. Though there is some doubt as to whether the full expression of phase status is to be found in every instance, the total observed D^2 contributed by each combination of characters is given in Table 1 for two of the species. Among the conventional characters, the head-width and pronotal length seem to be most useful with this species. In exploratory investigations one needs a battery of such characters in the hope that, between them, the several underlying modes of growth will be sufficiently illuminated.

DISCUSSION

There are two broad purposes served by the construction of generalised distance charts. That emphasised by Rao [1952] throws into relief the proximity of groups represented on the chart; those separated by a small distance are more alike morphometrically than are pairs with a high value between them. An extension follows from the recognition that discriminant functions are vectors which can differ in both magnitude (D^2-value) and direction as first noted by Fisher [1938].

Locusts and grasshoppers have a well-marked sexual dimorphism which is represented on a generalised distance chart by a series of nearly parallel vectors joining the sexes. By including such size-sensitive characters as body-weight at eclosion in the discriminant function and noting the increment of the D^2-statistic produced by each character relative to that produced by including the weight, the size-sensitivity of any character can be tested. *Mutatis mutandi*, the efficacy of any character as a discriminator of phases or species may be examined. The generalised distance charts for all the swarming species so far investigated are two-dimensional, confirming the qualitative distinction between modes of normal and of density-dependent growth.

Further comparisons between the species, wherever these have been made, show that the specific differences of form are not, in these instances, an exaggeration of the density-dependent or normal modes of growth. This kind of evidence, albeit negative in character, is of taxonomic interest, and is available at those lower taxonomic levels for which the conventional classificatory processes are often least securely based on a phylogenetic foundation. One could use such methods to help elucidate such current problems as the gradual, or, alternatively, the saltatory, nature of speciation.

To show that putative species differ little along the 'species dimension', is not to demonstrate their identity. One could do as much for many pairs of well-defined species. The present illustrations are included to show one kind of relevant information which may be extracted from multivariate analyses of form in locusts. Key [1954] figures a tentative phylogenetic tree of the locusts he studied, based on characters of taxonomic value, on ecology, and on distribution. The generalised distance charts may be regarded as a morphometric reflection of such a tree. The correspondence is encouraging. The genera are distinctly separated; *tenuicornis* is on both representations the *Austroicetes* closest in affinity to *Chortoicetes*; and *pusilla*, *cruciata* (Eastern race) and *nullarborensis* are the *Austroicetes* least like *Chortoicetes*. The most serious discrepancy is that for the Western race of *A. cruciata*. Besides being internally inconsistent, in that the *solitaria* show a sexual dimorphism along the 'species' dimension, all four groups of the Western race are far removed from those of the Eastern race, for which feature no easy explanation seems available. The general success of the technique adopted in disclosing biologically identifiable dimensions of variation, despite the small size of many samples, may yet help to illustrate a use for discriminatory analysis complementary to the current preoccupation with classificatory problems (Williams [1955]).

I should like to thank Dr. K. H. L. Key for reading this note in

manuscript and for many helpful suggestions, and Miss Margaret Roberts for most of the computations involved. I owe the photographs and the projection to my colleague Mr. J. W. Siddorn.

REFERENCES

Albrecht, F. O. [1955]. La densité des populations et la croissance chez *Schistocerca gregaria* (Forsk.) et *Nomadacris septemfasciata*, (Serv.): la mue d'adjustment. *Journ. agric. trop. bot. appl. 2*; 110–192.

Albrecht, F. O. and Blackith, R. E. [1957]. Phase and developmental polymorphism in locusts. *Evolution.*

Anderson, E. [1953]. Introgressive hybridization. *Biol. Revs., Cambridge. 28*, 280–307.

Burt, C. and Banks, C. [1947]. A factorial analysis of body measurements for British adult males. *Ann. Eugen. Camb. 13*, 238–256.

Clark, F. Le Gros and Medawar, P. B. [1945] (Eds.) *Essays on growth and form.* Clarendon Press, Oxford.

Cousin, G. [1948]. Hybridizations interspécifiques et transmissions des types structureaux. *Bull. Biol. Fr. Belg. 82*, 331–387.

Fisher, R. A. [1938]. The statistical utilization of multiple measurements. *Ann. Eugenics, Cambridge, 8*, 376–386.

Fraisse, R. and Arnoux, J. [1954]. Les caractères biométriques du cocon chez *Bombyx mori*, L. et leurs variations sur l'influence de l'alimentation. *Revue du Ver à Soie. 6*, 43–62.

Goodall, D. W. [1953]. Objective methods for the classification of vegetation 2, Fidelity and indicator value. *Aust. J. Bot. 1*, 434–456.

Hotelling, H. [1954]. Multivariate analysis, *Statistics and Mathematics in Biology.* Ed. Kempthorne, Bancroft, Gowen, Lush. Iowa State College Press. pp. 67–80.

Hughes, R. E. and Lindley, D. V. [1955]. Application of biometric methods to problems of classification in ecology. *Nature (Lond.) 175*, 806–807.

Huxley, J. S. [1932]. *Problems of relative growth.* London.

Kendall, M. G. [1946]. *Advanced theory of statistics* Vol. II. Charles Griffin and Company, Ltd., London, p. 341.

Key, K. H. L. [1954]. The taxonomy, phases, and distributions of the genera *Chortoicetes* Brunn. and *Austroicetes* Uv. (Orthoptera; Acrididae) Canberra, Commonwealth Scientific and Industrial Research Organization.

Key, K. H. L. and Day, M. F. [1954]. A temperature-controlled physiological colour response in the grasshopper *Kosciuscola tristis* Sjöst. (Orthoptera; Acrididae) *Aust. J. Zool. 2*, 309–339.

Mahalanobis, P. C., Majumdar, D. N. and Rao, C. R. [1949]. Anthropometric survey of the United Provinces, 1941. A statistical study. *Sankhyā. 9*, 89–360.

Mukherjee, R., Rao, C. R. and Trevor, J. C. [1955]. *The Ancient Inhabitants of the Jebel Moya.* Cambridge University Press, Appendix 4.

Pasquier, R. [1938]. La lutte rationelle contre le criquet marocain en Algéric. *C.r. 5 me. Conf. Int. Rech. anti-acrid.* (Bruxelles) 367–379.

Rao, C. R. [1952]. *Advanced statistical methods in biometric research.* Wiley, New York.

Teissier, G. [1937]. Essai d'analyse de la variabilité des *Maia squinado*. *C. R. Soc. Biol. (Paris) 124*, 1073–1075.

Teissier, G. [1948]. La relation d'allométrie: sa signification statistique et biologique. *Biometrics 4,* 14–53.

Teissier, G. [1955a]. Grandeur de référence et allométrie de taille chez *Maia squinado*. *C.R. Acad. Sciences. 240,* 364–366.

Teissier, G. [1955b]. Sur la détermination de l'axe d'un nuage rectiligne de points. *Biometrics. 11,* 344–356.

Uvarov, B. P. [1921]. A revision of the genus *Locusta* L. (= *Pachytylus* Fieb.) with a new theory as to the periodicity and migrations of locusts. *Bull. Ent. Res., Lond. 12,* 135–163.

White, M. J. D. and Key, K. H. L. [1957]. Acytotaxonomic study of the *Pusilla* group of species in the genus *Austroicetes* Uv. (Orthoptera; Acrididae) *Aust. J. Zool. 5,* 56–87.

Williams, E. J. [1955]. Significance tests for discriminant functions and lineal functional relationships. *Biometrika, 42,* 360–381.

EXAMPLES OF INTRA-BLOCK ANALYSIS FOR FACTORIALS IN GROUP DIVISIBLE, PARTIALLY BALANCED, INCOMPLETE BLOCK DESIGNS*,**

CLYDE YOUNG KRAMER AND RALPH ALLAN BRADLEY
Virginia Agricultural Experiment Station of the Virginia Polytechnic Institute, Blacksburg, Virginia, U.S.A.

1. INTRODUCTION

Incomplete block designs were developed to accommodate experiments wherein it is necessary that the number of experimental units per block be less than the number of treatments. These designs are used extensively in many fields of research. Factorial treatment combinations were introduced so that the effects of several variables, together with their interactions, may be studied in a single experiment. Factorial treatment combinations are being used in many fields of research. Kramer [1] and Kramer and Bradley [2] have presented the theory allowing one to combine these two important concepts, incomplete blocks and factorials, to increase the utility of both.

Partially balanced incomplete block designs were first discussed by Bose and Nair [8] and these designs have proved to be very useful. Bose and Shimamoto [9] later classified partially balanced incomplete block designs and, since then, Bose, Clatworthy, and Shrikhande [3] have provided a catalogue of such designs with two associate classes. Designs are given in the catalogue for block sizes $3 \leq k \leq 10$ and replications $r \leq 10$ and when E_1 and E_2, efficiency factors that will be further defined later, are not too different. Group divisible designs form a large and important class of partially balanced, incomplete block designs with two associate classes. In the catalogue, the parameters for each design, the association matrices, and block layouts are given. Clatworthy [4] has considered designs with $k = 2$ and $r \leq 10$

*Research sponsored by the Agricultural Research Service, U.S.D.A., under a Research and Marketing Act Contract, No. 12-14-100-126(20).

**A condensation of part of a dissertation by C. Y. Kramer submitted to the Virginia Polytechnic Institute in partial fulfillment of the requirements for the Ph.D. degree in Statistics.

and indicates how such designs may be easily written down. It is the objective of this paper to present examples of analyses for factorial treatments in group divisible, incomplete block designs and to summarize the basic general results obtained.

Factorials in incomplete block designs, except as confounding and partial confounding have been developed, seem not to have been much used. Methods have been available for the use of factorials in balanced incomplete block designs since 1938, the date of publication by Cornish [10]. In addition, Harshbarger [7] considered a 2^3 factorial in a Latinized, rectangular lattice. In current statistical practice, there appears to be an increasing need for factorials in incomplete block designs with applications in both industry and agriculture.

When this work was presented at a conference,* R. C. Bose noted that the designs in [3] were essentially developed for varietal trials and that this was why E_1 and E_2 were not too different. He pointed out that, for factorial treatment combinations, designs with widely different values of E_1 and E_2 become much more important, for it may be that much is already known about some factors and their interactions while others need to be more fully investigated. Then the factorial treatments may be assigned to the association matrix in accordance with these needs. Certain group divisible designs are *disconnected* and have been discarded and not catalogued on the grounds that they are not useful designs for varietal trials. These designs may be very useful for factorials, for then the factorial associations of the treatment combinations act as links between blocks.

In developing the theory for factorials, we have presented the analysis of variance with the treatment sum of squares as a function of the least squares estimators of the treatment effects. The sums of squares for factorial effects are also obtained simply in terms of these treatment estimators. We have thus somewhat reduced the number of parameters of the designs that are required and explicitly shown in the analyses. It will be seen that the analysis of variance, both for unrelated treatments and for factorial treatments, in group divisible designs is not more difficult than for simpler incomplete block design that are more frequently used.

The results in this paper are more general than those obtained by Cornish. Balanced incomplete block designs may be regarded as a special class of group divisible designs, and we have noted the special forms obtained for balanced designs as developed by Cornish and included here for the convenience of the reader.

*Industrial Experimental Design Conference sponsored by the Air Force Office of Scientific Research, N. C. State College, November 5–9, 1956.

2. GROUP DIVISIBLE, PARTIALLY BALANCED, INCOMPLETE BLOCK DESIGNS

The properties of group divisible designs given in [3] are repeated here for convenience. These properties are:

(i) The design has b blocks of k experimental units with different treatments (or treatment combinations) applied to the units in the same block.

(ii) There are $v = mn$ treatments ($v > k$) assignable to m groups of n each in the association scheme (1), where we use a double subscript notation instead of the more usual single subscripts, such that treatments in the same group or row are first associates and two treatments not in the same row of (1) are second associates.

$$\begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ V_{m1} & V_{m2} & \cdots & V_{mn} \end{bmatrix} \tag{1}$$

(iii) Each treatment has $(n - 1)$ first associates and $n(m - 1)$ second associates.

(iv) Two treatments that are i-th associates appear together in exactly λ_i blocks, $i = 1, 2$.

(v) Given any two treatments that are i-th associates, the number of treatments common to the j-th associate of the first and the k-th associate of the second is p^i_{jk}, and is independent of the pair of treatments selected. If P_i is the matrix with elements p^i_{jk},

$$P_1 = \begin{bmatrix} (n-2) & 0 \\ 0 & n(m-1) \end{bmatrix} \quad \text{and} \quad P_2 = \begin{bmatrix} 0 & (n-1) \\ (n-1) & n(m-2) \end{bmatrix}.$$

(vi) The inequalities, $r \geq \lambda_1$, $rk - \lambda_2 v \geq 0$, hold.

(vii) The design parameters are related so that $(n - 1)\lambda_1 + n(m - 1)\lambda_2 = r(k - 1)$ or $rk - \lambda_2 v = r - \lambda_1 + n(\lambda_1 - \lambda_2)$.

Subclasses of these designs are:

Singular (S) if $r = \lambda_1$,
Semi-regular (SR) if $r > \lambda_1$ and $rk - \lambda_2 v = 0$, and
Regular (R) if $r > \lambda_1$ and $rk - \lambda_2 v > 0$.

We shall consider the class as a whole without subdivision. But first let us note that certain designs, familiar to the reader and described by Harshbarger [5, 6], are a subclass of the Semi-regular designs.*

*The grouping of blocks to form near balance, rectangular lattices or Latinized, rectangular lattices only requires subdivision of the unadjusted block sum of squares (as in Table 1).

These are near balance, rectangular lattices and Latinized, rectangular lattices, and they may be defined in terms of integers K and Q associated with the parameters of the semi-regular designs as follows: $v = K(K - Q)$, $k = K - Q$, $r = K$, $b = K^2$, $m = K - Q$, $n = K$, $\lambda_1 = 0$, and $\lambda_2 = 1$.

When $\lambda_1 = \lambda_2$, the incomplete block design is balanced. The results of this paper apply to balanced incomplete block designs and simplify through use of the stated equality. Then treatments may be assigned to an association matrix like (1) in any desired way and with free choice of values of m and n such that $v = mn$.

The model assumed for all randomized incomplete block designs is

$$y_{ijs} = \mu + \tau_{ij} + \beta_s + \epsilon_{ijs} , \qquad (2)$$

where y_{ijs} is the observation on V_{ij} in block s if that treatment occurs in block s, μ is the grand mean, τ_{ij} is the effect of V_{ij}, β_s is the effect of block s, and ϵ_{ijs} is the usual normal random error with mean zero and variance σ^2, the various ϵ's being independent. Restrictions on the parameters in (2) are that $\sum_{i=1}^{m} \sum_{j=1}^{n} \tau_{ij} = 0$ and $\sum_{s=1}^{b} \beta_s = 0$. The analysis of variance is given algebraically (with the exception of mean squares which are obtained in the usual way from sums of squares) in Table 1. The meaning of Table 1 is clear with the following definitions:

TABLE 1

ANALYSIS OF VARIANCE FOR THE GENERAL MODEL

Source	D.f.	Sum of squares	Ref. No.
Treatments (adj.)	$v - 1$	$\frac{(\lambda_1 + rk - r)}{k} \sum_{i=1}^{m} \sum_{j=1}^{n} t_{ij}^2 + \frac{(\lambda_2 - \lambda_1)}{k} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} t_{ij} \right)^2$	(3)
Blocks (unadj.)	$b - 1$	$\frac{1}{k} \sum_{s=1}^{b} B_s^2 - \frac{G^2}{rv}$	(4)
Intra-block error	$[v(r-1) - b + 1]$	By subtraction	
Total	$rv - 1$	$\sum_{s=1}^{b} \sum_{\substack{i \\ \text{in} \\ s}} \sum_{j} y_{ijs}^2 - \frac{G^2}{rv}$	(5)

$$G = \sum_{s=1}^{b} \sum_{\substack{i \text{ in} \\ s}} \sum_{j} y_{iis} , \tag{6}$$

$$B_s = \sum_{\substack{i \\ \text{in} \\ s}} \sum_{j} y_{iis} , \tag{7}$$

$$t_{ij} = [kv\lambda_2 T_{ij} - k(\lambda_2 - \lambda_1) \sum_{h=1}^{n} T_{ih} - v\lambda_2 B_{ij.} + (\lambda_2 - \lambda_1) \sum_{h=1}^{r} B_{ih.}]/v\lambda_2(\lambda_1 + rk - r) \tag{8}$$

and t_{ij} is the least squares estimator* of τ_{ij} in (2),

$$T_{ij} = \sum_{\substack{s \\ \text{with} \\ (i,j)}} y_{iis} , \tag{9}$$

and

$$B_{ij.} = \sum_{\substack{s \\ \text{with} \\ (i,j)}} B_s . \tag{10}$$

---

*For a near balanced or Latinized design,

$$t_{ij} = [K(K-Q)^2 T'_{ij} - (K-Q) \sum_{h=1}^{K} T_{ih} - K(K-Q)B_{ij.} + \sum_{h=1}^{K} B_{ih.}]/K^2(K-Q)(K-Q-1)$$

$$= \frac{K(K-Q)-1}{K^2(K-Q-1)} T_{ij} - \sum_{\substack{h=1 \\ h \neq j}}^{K} \frac{T_{ih}}{K^2(K-Q-1)} - \frac{B_{ij.}}{K(K-Q-1)} + \frac{G}{K^2(K-Q)(K-Q-1)} ,$$

the latter result being in the usual textbook form. Now (3) becomes

$$\frac{K(K-Q-1)}{K-Q} \sum_{i=1}^{K-Q} \sum_{j=1}^{K} t_{ij}^2 + \frac{1}{K-Q} \sum_{i=1}^{K-Q} \left(\sum_{j=1}^{K} t_{ij} \right)^2 .$$

For a balanced design,

$$t_{ij} = (kT_{ij} - B_{ij.})/(\lambda + rk - r) = (kT_{ij} - B_{ij.})/v\lambda$$

and (3) becomes

$$\frac{v\lambda}{k} \sum_{i=1}^{m} \sum_{j=1}^{n} t_{ij}^2 ,$$

since then $\lambda + rk - r = v\lambda$.

More simply, G is the grand total, B_s is the block total for block s, T_{ij} is the treatment total for V_{ij}, and $B_{ij.}$ is the total of block totals for blocks containing V_{ij}. Convenient computing organization for obtaining the entries in Table 1 is indicated in the numerical examples.

Certain variances and covariances are useful. In the following results, s^2, the error mean square obtainable from Table 1, is used as the estimator of σ^2. We note that

$$V(t_{ij}) = k\sigma^2\left[\frac{(n-1)}{n(\lambda_1 + rk - r)} + \frac{(m-1)}{mn\lambda_2 v}\right], \tag{11}$$

$$\text{Cov}\,(t_{ij}t_{ij'}) = k\sigma^2\left[\frac{(m-1)}{mn\lambda_2 v} - \frac{1}{n(\lambda_1 + rk - r)}\right], \qquad j \neq j', \tag{12}$$

and

$$\text{Cov}\,(t_{ij}t_{i'j'}) = -k\sigma^2/mn\lambda_2 v, \qquad i \neq i'. \tag{13}$$

V_{ij} and $V_{i'j'}$ are first-associate treatments if and only if $i = i'$. Then,

$$V(t_{ij} - t_{ij'}) = 2k\sigma^2/(\lambda_1 + rk - r), \qquad j \neq j', \tag{14}$$

and

$$V(t_{ij} - t_{i'j'}) = 2k\sigma^2(\lambda_1 + \lambda_2 v - \lambda_2)/v\lambda_2(\lambda_1 + rk - r), \qquad i \neq i'. \tag{15}$$

The efficiencies E_1 and E_2, noted in the Introduction, depend on (14) and (15). These efficiencies are obtained by taking the ratio of the variance of the treatment contrast for a randomized block design to the corresponding variance for the incomplete block design, given equal values of r and on the assumption that both designs yield the same experimental error (that is, on the assumption that the use of the smaller blocks was not effective). The efficiency for the comparison of two treatments that are first associates is

$$E_1 = (\lambda_1 + rk - r)/rk \tag{16}$$

and that are second associates is

$$E_2 = v\lambda_2(\lambda_1 + rk - r)/rk(\lambda_1 + \lambda_2 v - \lambda_2). \tag{17}$$

The theory for the basic analysis of Table 1, and for the analyses for factorials summarized in the next section, may be developed by use of the method of least squares. To indicate the procedure, we note that to obtain the adjusted treatment sum of squares in Table 1, it is sufficient to evaluate the difference between the minima of the sums of

squares,

$$\sum_{s=1}^{b} \sum_{i} \sum_{\substack{j \\ \text{in } s}} (y_{ijs} - \mu - \tau_{ij} - \beta_s)^2$$

and

$$\sum_{s=1}^{b} \sum_{i} \sum_{\substack{j \\ \text{in } s}} (y_{ijs} - \mu - \beta_s)^2,$$

the former minimized subject to the restraints on the τ_{ij} and the β_s and the latter subject to the restraint on the β_s. Similar differences are required for sums of squares assignable to factorial effects and detail on the theory is given in [2].

3. FACTORIALS IN GROUP DIVISIBLE DESIGNS

We first consider what we regard as the basic two-factor factorial in group divisible, partially balanced, incomplete block designs. The analyses for multi-factor factorials follow from this.

Consider factors A and C with m and n levels respectively. We amend the model (2) to obtain

$$y_{ijs} = \mu + \alpha_i + \gamma_j + \delta_{ij} + \beta_s + \epsilon_{ijs} \tag{18}$$

with the restrictions,

$$\sum_{i=1}^{m} \alpha_i = 0, \qquad \sum_{j=1}^{n} \gamma_j = 0, \qquad \sum_{i=1}^{m} \delta_{ij} = 0,$$

$$\sum_{j=1}^{n} \delta_{ij} = 0, \quad \text{and} \quad \sum_{s=1}^{b} \beta_s = 0.$$

α_i is the effect of the i-th level of the A-factor, γ_j is the effect of the j-th level of the C-factor, and δ_{ij} represents the interaction of the i-th level of A with the j-th level of C. ϵ_{ijs} is as previously defined. The new model (18) is obtained by setting τ_{ij} in (2) equal to $(\alpha_i + \gamma_j + \delta_{ij})$. The treatment V_{ij} has now become the factorial treatment combination A_iC_j and such substitution may be made in the association matrix (1) if desired.

The least squares estimators for the factorial effects in (18) may be expressed in terms of the original treatment estimators t_{ij} in (8). We now have, using latin letters for the corresponding parameters,

$$a_i = \frac{1}{n} \sum_{j=1}^{n} t_{ij} = \bar{t}_{i\cdot}, \tag{19}$$

$$c_j = \frac{1}{m} \sum_{i=1}^{m} t_{ij} = \bar{t}_{\cdot j}, \tag{20}$$

and

$$d_{ij} = t_{ij} - \bar{t}_{i.} - \bar{t}_{.j} \,. \tag{21}$$

The appropriate analysis of variance is given in Table 2. In Table 2, the sums of squares (22), (23), and (24) add to the treatment sum of squares (3) in Table 1, and consequently it is easier to obtain the adjusted AC-interaction sum of squares by subtracting the total of (22) and (23) from (3).* Sums of squares for factorial effects in Table

TABLE 2

ANALYSIS OF VARIANCE FOR THE TWO-FACTOR FACTORIAL

Source	D.f.	Sum of squares	Ref. no.
A-factor (adj.)	$m - 1$	$\frac{n\lambda_2 v}{k} \sum_{i=1}^{m} \bar{t}_{i.}^2$	(22)
C-factor (adj.)	$n - 1$	$\frac{m(\lambda_1 + rk - r)}{k} \sum_{j=1}^{n} \bar{t}_{.j}^2$	(23)
AC-interaction (adj.)	$(m - 1)(n - 1)$	$\frac{(\lambda_1 + rk - r)}{k} \cdot \sum_{i=1}^{m} \sum_{j=1}^{n} (t_{ij} - \bar{t}_{i.} - \bar{t}_{.j})^2$	(24)
Blocks (unadj.)	$b - 1$	$\frac{1}{k} \sum_{s=1}^{b} B_s^2 - \frac{G^2}{rv}$	
Intra-block error	$[v(r - 1) - b + 1]$	By subtraction	
Total	$rv - 1$	$\sum_{s=1}^{b} \sum_{\substack{i \\ \text{in} \\ s}} \sum_{j} y_{ijs}^2 - \frac{G^2}{rv}$	

*For a near balance or Latinized design, (22), (23), and (24) become respectively

$$K^2 \sum_{i=1}^{K-Q} \bar{t}_{i.}^2 \,, \qquad K(K - Q - 1) \sum_{j=1}^{K} \bar{t}_{.j}^2 \,,$$

$$\text{and} \quad \frac{K(K - Q - 1)}{(K - Q)} \sum_{i=1}^{K-Q} \sum_{j=1}^{K} (t_{ij} - \bar{t}_{i.} - \bar{t}_{.j})^2 .$$

For a balanced design, they become

$$\frac{n\lambda v}{k} \sum_{i=1}^{m} \bar{t}_{i.}^2 \,, \qquad \frac{m\lambda v}{k} \sum_{j=1}^{n} \bar{t}_{.j}^2 \,, \quad \text{and} \quad \frac{\lambda v}{k} \sum_{i=1}^{m} \sum_{j=1}^{n} (t_{ij} - \bar{t}_{i.} - \bar{t}_{.j})^2 .$$

In a group divisible, partially balanced design, m and n are specified by the design; in a balanced design, m and n may be chosen in any way so long as $v = mn$.

2 are independent. (22), (23), and (24) respectively are used for tests of the null hypotheses,

$$H_0(A): \alpha_1 = \cdots = \alpha_m ,$$

$$H_0(C): \gamma_1 = \cdots = \gamma_n , \quad \text{and}$$

$$H_0(AC): \delta_{11} = \cdots = \delta_{mn} ,$$

against general alternatives in each case.

The following variances are estimated by substituting the error mean square obtainable from Table 2 for σ^2 .

$$V(a_i - a_{i'}) = \frac{2k\sigma^2}{n\lambda_2 v} , \qquad i \neq i', \tag{25}$$

$$V(c_j - c_{j'}) = \frac{2k\sigma^2}{m(\lambda_1 + rk - r)} , \qquad j \neq j', \tag{26}$$

and

$$\begin{aligned} V(d_{ij}) &= \frac{(m-1)(n-1)k\sigma^2}{mn(\lambda_1 + rk - r)} \\ \text{Cov}\,(d_{ij}\, d_{i'j}) &= \frac{-k(n-1)\sigma^2}{mn(\lambda_1 + rk - r)} , \qquad i \neq i', \\ \text{Cov}\,(d_{ij}\, d_{ij'}) &= \frac{-k(m-1)\sigma^2}{mn(\lambda_1 + rk - r)} , \qquad j \neq j', \\ \text{Cov}\,(d_{ij}\, d_{i'j'}) &= \frac{k\sigma^2}{mn(\lambda_1 + rk - r)} , \qquad i \neq i', \qquad j \neq j'. \end{aligned} \tag{27}$$

Efficiencies for factorial contrasts, similar to E_1 and E_2 in (16) and (17), are

$$E_A = \lambda_2 v/rk \tag{28}$$

for an A-factor contrast,

$$E_C = (\lambda_1 + rk - r)/rk \tag{29}$$

for a C-factor contrast, and

$$E_{AC} = (\lambda_1 + rk - r)/rk \tag{30}$$

for an AC-interaction contrast. Note that, for the singular subclass of two-associate class, group divisible designs, $E_C = E_{AC} = 1$, while, for the semi-regular subclass, $E_A = 1$. For a balanced design, $E_A = E_C = E_{AC} = E$, the efficiency of the design.

Single degree of freedom contrasts may be obtained in much the usual way. To partition the effects of the m-level A-factor into $(m - 1)$ individual contrasts, each yielding an adjusted sum of squares with one degree of freedom, we have in effect only to make an orthogonal transformation on the m estimators, $\bar{t}_{1.}, \cdots, \bar{t}_{m.}$. Let the u-th such contrast be

$$I_u = \sum_{i=1}^{m} \xi_{iu} \bar{t}_{i.}, \qquad u = 1, \cdots, (m - 1). \tag{31}$$

(In the practice of analysis of variance, $\xi_{1u}, \cdots, \xi_{mn}$, will be a set of real coefficients that sum to zero.) The adjusted sum of squares for a test of the hypothesis that $\sum_{i=1}^{m} \xi_{iu}\alpha_i = 0$ is

$$\begin{aligned} \text{Adj. S. S. } (I_u) &= \frac{n\lambda_2 v}{k \sum_{i=1}^{m} \xi_{iu}^2} \left(\sum_{i=1}^{m} \xi_{iu} \bar{t}_{i.} \right)^2 \\ &= \frac{\lambda_2 v}{k \sum_{i=1}^{m} \sum_{j=1}^{n} \xi_{iu}^2} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} \xi_{iu} t_{ij} \right)^2. \end{aligned} \tag{32}$$

In the same way, let the v-th contrast among C-factor effects be

$$J_v = \sum_{j=1}^{n} \eta_{jv} \bar{t}_{.j}, \qquad v = 1, \cdots, (n - 1) \tag{33}$$

and the adjusted sum of squares for a test of the hypothesis that $\sum_{j=1}^{n} \eta_{jv}\gamma_j = 0$ is

$$\begin{aligned} \text{Adj. S. S. } (J_v) &= \frac{m(\lambda_1 + rk - r)}{k \sum_{j=1}^{n} \eta_{jv}^2} \left(\sum_{j=1}^{n} \eta_{jv} \bar{t}_{.j} \right)^2 \\ &= \frac{(\lambda_1 + rk - r)}{k \sum_{i=1}^{m} \sum_{j=1}^{n} \eta_{jv}^2} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} \eta_{jv} t_{ij} \right)^2. \end{aligned} \tag{34}$$

The adjusted interaction sum of squares may also be partitioned. The contrast for interaction of I_u and J_v is

$$\text{Adj. S. S. } (I_u J_v) = \frac{(\lambda_1 + rk - r)}{k \sum_{i=1}^{m} \sum_{j=1}^{n} (\xi_{iu}\eta_{jv})^2} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} \xi_{iu}\eta_{jv} t_{ij} \right)^2. \tag{35}$$

The way is now open to consider multi-factor factorials. Suppose the levels of the A-factor are themselves factorial combinations of g

factors, $A^{(1)}, \cdots, A^{(g)}$, with levels $m_1, \cdots, m_g$, $m = \prod_{x=1}^{g} m_x$. Then appropriate choice of contrasts I_u, $u = 1, \cdots, (m-1)$, permits computation of adjusted sums of squares for all of the main effects and interactions among $A^{(1)}, \cdots, A^{(g)}$. In the same way, the levels of the C-factor may be factorial treatment combinations of factors $C^{(1)}, \cdots, C^{(h)}$ with levels $n_1, \cdots, n_h$, $n = \prod_{w=1}^{h} n_w$. Then appropriate choice of contrasts J_v, $v = 1, \cdots, (n-1)$, permits computation of the adjusted sums of squares for these factorial effects. The adjusted interaction sums of squares, Adj. S. S. $(I_u J_v)$, in (34) may be identified with interactions of main effects and interactions of $A^{(1)}, \cdots, A^{(g)}$ with main effects and interactions of $C^{(1)}, \cdots, C^{(h)}$. The computations for multi-factor factorials will be illustrated in the examples that follow.

Note also that the levels of the A-factor may be associated with the treatment combinations of a fraction of a multi-factor factorial and so for the levels of the C-factor. Again appropriate choice of I_u and J_v leads easily to the required analysis. This also will be illustrated in one of the examples.

The efficiency factor E_A applies to all main-effect and interaction contrasts among $A^{(1)}, \cdots, A^{(g)}$, E_C to those among $C^{(1)}, \cdots, C^{(h)}$, and E_{AC} to interactions between A-factor factorial effects and C-factor factorial effects.

4. FIRST EXAMPLE, A GROUP DIVISIBLE DESIGN

We now illustrate the applications of the methods outlined. Two examples will be given, one on a near balance design and one on a group divisible design that is not a near balance design. It was necessary at the time of writing to use made-up examples but they indicate how the methods may be used.

(*i*) *Basic Analysis*

We first consider use of the group divisible design $S6$ in [3] with design parameters,

$$v = 8, \quad r = 3, \quad k = 4, \quad b = 6, \quad m = 4, \quad n = 2, \quad \lambda_1 = 3, \quad \lambda_2 = 1,$$

and do the basic analysis of variance without regard to factorial make-up of treatments. The association matrix (1) now has four rows and two columns. Block lay-outs, observations, block totals, and grand total are shown in Table 3. To compute the t_{ij} in (8), it is convenient to further summarize the data as in Table 4. We require T_{ij}, $\sum_{j=1}^{2} T_{ij}$, $B_{ij.}$, and $\sum_{j=1}^{2} B_{ij.}$ and these may be obtained easily from Table 3; one may split the keyboard on a calculating machine, add

TABLE 3

OBSERVATIONS AND TOTALS FOR DESIGN *S*6

V_{11} 29	V_{12} 38	V_{21} 40	V_{22} 33	B_1 140
V_{31} 28	V_{32} 37	V_{41} 24	V_{42} 24	B_2 113
V_{11} 20	V_{12} 37	V_{31} 26	V_{32} 33	B_3 116
V_{21} 23	V_{22} 22	V_{41} 24	V_{42} 15	B_4 84
V_{11} 27	V_{12} 41	V_{41} 36	V_{42} 28	B_5 132
V_{21} 37	V_{22} 26	V_{31} 29	V_{32} 37	B_6 129
			Total	G 714

TABLE 4

VALUES OF T_{ij}, $B_{ij.}$, $\sum_i T_{ij}$, $\sum_i B_{ij.}$, t_{ij}, $t_{i.}$, $t_{.j}$, $\bar{t}_{i.}$, AND $\bar{t}_{.j}$ FOR THE FIRST EXAMPLE, DESIGN *S*6

	T_{ij}		$\sum_{j=1}^{2} T_{ij}$	$B_{ij.}$		$\sum_{j=1}^{2} B_{ij.}$	t_{ij}		$t_{i.}$	$\bar{t}_{i.}$
Values i	Values of j 1	2		Values of j 1	2		Values of j 1	2		
1	76	116	192	388	388	776	−7.167	6.167	−1.000	−0.500
2	100	81	181	353	353	706	4.292	−2.042	2.250	1.125
3	83	107	190	358	358	716	−1.250	6.750	5.500	2.750
4	84	67	151	329	329	658	−0.542	−6.208	−6.750	−3.375
						Values of $t_{.j}$	−4.667	4.667		
						Values of $\bar{t}_{.j}$	−1.167	1.167		

observations on one side to obtain T_{ij} , and add corresponding block totals on the other side to obtain $B_{ij.}$. In the same way, $\sum_{j=1}^{2} T_{ij}$ and $\sum_{j=1}^{2} B_{ij.}$ may be obtained in one operation. We have also included values t_{ij} , $t_{i.}$, $t_{.j}$, $\bar{t}_{i.}$, and $\bar{t}_{.j}$ in Table 4, for this is a convenient place to summarize these calculated values. Values of t_{ij} are computed from (8) and, for design *S*6, (8) becomes

$$t_{ij} = \frac{1}{3} T_{ij} + \frac{1}{12} \sum_{h=1}^{2} T_{ih} - \frac{1}{12} B_{ij.} - \frac{1}{48} \sum_{h=1}^{2} B_{ih.} .$$

To illustrate how the entries in Table 4 are obtained, we note that

$$T_{11} = (29 + 20 + 27) = 76, \qquad B_{11.} = (140 + 116 + 132) = 388,$$

$$\sum_{j=1}^{2} T_{1j} = (76 + 116) = 192, \qquad \sum_{j=1}^{2} B_{1j.} = (388 + 388) = 776,$$

$$t_{11} = \frac{1}{3}(76) + \frac{1}{12}(192) - \frac{1}{12}(388) - \frac{1}{48}(776) = -7.167.$$

Recall that

$$t_{i.} = \sum_{j=1}^{2} t_{ij} , \qquad t_{.j} = \sum_{i=1}^{4} t_{ij} , \qquad \bar{t}_{i.} = \frac{1}{2} t_{i.} , \qquad \bar{t}_{.j} = \frac{1}{4} t_{.j} .$$

The sums of squares for the analysis of variance are now computed from (3), (4), and (5) in Table 1. Total and unadjusted block sums of squares are computed in the usual way. The adjusted treatment sum of squares (3) becomes for design *S*6

$$\text{Adj. Treat. S. S.} = 3 \sum_{i=1}^{4} \sum_{j=1}^{2} t_{ij}^2 - \frac{1}{2} \sum_{i=1}^{4} \left(\sum_{j=1}^{2} t_{ij} \right)^2 = 552.88.$$

The error sum of squares is obtained by subtraction, mean squares are obtained in the usual way, and the complete analysis of variance is given in Table 5 corresponding to the general Table 1.

TABLE 5

BASIC ANALYSIS OF VARIANCE, DESIGN *S*6

Source	D.f.	S.s.	M.s.	F
Treatments (adj.)	7	552.88	79.98	10.52
Blocks (unadj.)	5	495.00	.....	.....
Intra-block error	11	82.62	7.51	.....
Total	23	1130.50	.....	.....

(*ii*) *Analysis for 4 by 2 Factorial*

We shall now suppose that the eight treatments resulted from a 4×2 factorial associated with the treatment designations so that $V_{ij} = A_iC_j$, A_i representing the i-th level of A, $i = 1, \cdots, 4$, and C_j representing the j-th level of C, $j = 1, 2$.

Adjusted sums of squares for factorial effects are obtained from (22) and (23) in Table 2. We have

$$\text{Adj. S. S. } (A) = 4 \sum_{i=1}^{4} \bar{l}_{i.}^2 = 81.87$$

and

$$\text{Adj. S. S. } (C) = 12 \sum_{j=1}^{2} \bar{l}_{.j}^2 = 32.67.$$

By subtraction of these two sums of squares from the adjusted treatment sum of squares in Table 5, we also obtain

$$\text{Adj. S. S. } (AC) = 552.88 - (81.87 + 32.67) = 438.34.$$

Alternate computation of the adjusted AC-interaction sum of squares may be used as a computing check and would be based on (24) in Table 2.

We shall take the factor A to be a quantitative one and subdivide the adjusted sum of squares for A into linear, quadratic, and cubic components. This is done using the first form of (32). The trend coefficients together with the sums of squares of the coefficients are given in Table 6. To illustrate, the adjusted sum of squares for the

TABLE 6

TREND COEFFICIENTS FOR SUBDIVISION OF ADJ. S. S. (A)

Contrasts	Coefficients for $\bar{l}_{1.} = -0.500$	$\bar{l}_{2.} = 1.125$	$\bar{l}_{3.} = 2.750$	$\bar{l}_{4.} = -3.375$	Sums of squared coefficients
Linear A	-3	-1	$+1$	$+3$	20
Quadratic A	$+1$	-1	-1	$+1$	4
Cubic A	-1	$+3$	-3	$+1$	20

linear A-component is

$$\text{Adj. S. S. (Linear } A) = \frac{(2)(1)(8)}{4(20)} \cdot [(-3)(-0.500) + \cdots + (3)(-3.375)]^2 = 9.80.$$

Similarly,

$$\text{Adj. S. S. (Quad. } A) = 60.06$$

and

$$\text{Adj. S. S. (Cubic } A) = 12.01.$$

The AC-interaction sum of squares could also be subdivided if we desired; this has not been done.

The analysis of variance for the 4 by 2 factorial is given in Table 7.

Referring back to (25) and (26), we note that

$$V(a_i - a_{i'}) = \frac{(2)(4)}{(2)(1)(8)}(7.51) = 3.75, \quad i \neq i', \quad i, \quad i' = 1, \cdots, 4,$$

and

$$V(c_1 - c_2) = \frac{(2)(4)}{(4)(3 + 12 - 3)}(7.51) = 1.25.$$

Variances and covariances of a_i or c_j may be quickly obtained, if desired, from use of (11), (12), and (13) and the definitions of a_i and c_j in (19) and (20).

(*iii*) *Analysis for* 2^3 *Factorial*

Suppose that the 4 by 2 factorial of the preceding subsection is really a 2^3 factorial by taking the four levels of A to be made up of two levels of a factor N and two levels of a factor P. The complete association of the treatment combinations of the 2^3-factorial with the treatments V_{ij} is given in Table 8. We note only that in this new situation we require a new subdivision of the adjusted sum of squares for A and a subdivision of the AC-interaction of the last subsection.

The contrasts required, each with one degree of freedom, are obtainable most easily from the second forms in (32) and (34) and from (35). The required orthogonal sets of coefficients are shown in Table 8 and the resultant analysis of variance in Table 9. We simply illustrate the computing by considering the adjusted sum of squares for NPC-interaction. Here we have

$$\text{Adj. S. S. } (NPC) = \frac{(3 + 12 - 3)}{(4)(8)}$$

$$\cdot[(-1)(-7.167) + (1)(6.167) + \cdots + (1)(-6.208)]^2 = 13.50.$$

(*iv*) *Analysis for a Fraction (one-half) of a* 2^4 *Factorial*

Let us think of the eight treatment combinations of Subsection

TABLE 7

ANALYSIS OF VARIANCE FOR THE 4 BY 2 FACTORIAL

Source	D.f.	S.s.	M.s.	F
Treatments (adj.)	7	552.88	78.98	10.52
A-Factor	3	81.87	27.29	3.63
Linear *A*	1	9.80	9.80	1.30
Quad. *A*	1	60.06	60.06	8.00
Cubic *A*	1	12.01	12.01	1.60
C-Factor	1	32.67	32.67	4.35
AC-Interaction	3	438.34	146.11	19.46
Blocks (Unadj.)	5	495.00	....	....
Intra-block error	11	82.62	7.51	....
Total	23	1130.50	....	....

TABLE 8

COEFFICIENTS FOR ANALYSIS AS A 2^3 FACTORIAL AND AS A HALF-FRACTION OF A 2^4 FACTORIAL

Treatments	V_{11}	V_{12}	V_{21}	V_{22}	V_{31}	V_{32}	V_{41}	V_{42}	
2^3 Factorial designations	$N_1P_1C_1$	$N_1P_1C_2$	$N_1P_2C_1$	$N_1P_2C_2$	$N_2P_1C_1$	$N_2P_1C_2$	$N_2P_2C_1$	$N_2P_2C_2$	
2^4 Factorial designations*	$N_1P_1C_1D_1$	$N_1P_1C_2D_2$	$N_1P_2C_1D_2$	$N_1P_2C_2D_1$	$N_2P_1C_1D_2$	$N_2P_1C_2D_1$	$N_2P_2C_1D_1$	$N_2P_2C_2D_2$	Sums of squared coefficients
Treatment estimators	$t_{11} =$ -7.167	$t_{12} =$ 6.167	$t_{21} =$ 4.292	$t_{22} =$ -2.042	$t_{31} =$ -1.250	$t_{32} =$ 6.750	$t_{41} =$ -0.542	$t_{42} =$ -6.208	
Contrasts									
N (PCD)	-1	-1	-1	-1	$+1$	$+1$	$+1$	$+1$	8
P (NCD)	-1	-1	$+1$	$+1$	-1	-1	$+1$	$+1$	8
C (NPD)	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	8
NP (CD)	$+1$	$+1$	-1	-1	-1	-1	$+1$	$+1$	8
NC (PD)	$+1$	-1	$+1$	-1	-1	$+1$	-1	$+1$	8
PC (ND)	$+1$	-1	-1	$+1$	$+1$	-1	-1	$+1$	8
NPC (D)	-1	$+1$	$+1$	-1	$+1$	-1	-1	$+1$	8

*These 2^4 factorial designations are discussed in Subsection (iv) and the contrasts in parentheses also refer to the discussion in (iv).

(iii) as a half fraction of a 2^4 factorial. We introduce the new factor D, with two levels, and use the defining contrast,

$$I = NPCD.$$

We can now associate the four-factor treatment combinations with the treatments V_{ij} as indicated at the top of Table 8. The analysis of variance now obtained is identical with that of Table 9 but we note that confounding is present. The factorial contrasts of the 2^3 factorial have as aliases additional contrasts of the 2^4 factorial. These additional contrasts are shown in parentheses in Table 9.

TABLE 9

ANALYSIS OF VARIANCE AS A 2^3 FACTORIAL AND AS A HALF-FRACTION OF A 2^4 FACTORIAL

Source	D.f.	S.s.	M.s.	F
Treatments (adj.)	7	552.88	78.98	10.52
*N (PCD)**	1	1.56	1.56	.....
P (NCD)	1	20.25	20.25	2.70
C (NPD)	1	32.67	32.67	4.35
NP (CD)	1	60.06	60.06	8.00
NC (PD)	1	8.17	8.17	1.09
PC (ND)	1	416.67	416.67	55.48
NPC (D)	1	13.50	13.50	1.80
Blocks (unadj.)	5	495.00	....	....
Intra-block error	11	82.62	7.51	....
Total	23	1130.50	....	....

*Contrasts in parentheses are aliases of the stated contrasts when we have the half-fraction of the 2^4 factorial discussed in Subsection (iv).

This fractional factorial is included to illustrate how they may be used in the incomplete block designs. The analysis of variance would be useful if, for example, we know in advance that the D-factor does not interact with the other three factors in the experiment. Interpretive difficulty would not enter except perhaps in the case of the D-factor main effect which is confounded with NPC-interaction. Interpretation of this contrast is easy if NPC-interaction may be assumed inconsequential; if this assumption cannot be made, the design can hardly be regarded as appropriate for the study of the additional factor D. We then should have selected a different defining contrast or recognized the necessity of using a complete factorial. Difficulties

in selecting the appropriate fraction of a multi-factor factorial for use in incomplete block designs are the same as those met in the use of fractional factorials in general and we shall not discuss them here.

5. SECOND EXAMPLE, A NEAR BALANCE DESIGN

In certain areas of application of experimental designs, lattice designs of various kinds have most frequently been used. Thus near balance, rectangular lattices and Latinized, rectangular lattices, while they fall in the more general class of group divisible designs, are better known than other designs of the general class. We have accordingly prepared an example that may be used to illustrate the use of factorials in the near balance, and Latinized, rectangular lattices and have already noted the special forms of formulas for near balance designs in general. A balanced lattice design is, of course, a balanced incomplete block design and the special formulas noted for balanced designs would apply.

To use catalogued or derived near balance, and Latinized, rectangular designs, it is quite easy to write down the required association matrix by looking at the design itself. Alternatively, these designs are also catalogued in [3] together with their association matrices. They are listed among the semi-regular designs but are not designated as lattice designs and blocks are not necessarily arranged in replications.

We think of a near balance design as a design with parameters specified in terms of K and Q as noted in Section 2. When the blocks are grouped into replications, the design becomes a near balance, rectangular lattice; when the blocks are grouped in a two-way pattern of rows and columns with each treatment in each row and column, the design becomes a Latinized, rectangular lattice. Our example is a four by three near balance design and is listed in [3] as design $SR21$. In this example $K = 4$, $K - Q = 3$, and $Q = 1$. Then $v = 12$, $r = 4$, $k = 3$, $b = 16$, $m = 3$, $n = 4$, $\lambda_1 = 0$, $\lambda_2 = 1$. We show the plot layouts, the observations, the block totals, and the grand total for our illustrative example in Table 10.

(*i*) *Basic Analysis*

Table 11 contains the treatment totals $T_{.i}$, the totals of block totals for blocks containing a specified treatment $B_{ij.}$, sums of these quantities, and values of the estimators t_{ij} with certain of their sums and averages. Table 11 is comparable to Table 4 in the first example. For this example, the association scheme (1) has three rows and four columns. T_{ij} and $B_{ij.}$ are obtained from Table 10 as before; the estimates t_{ij} may again be obtained from (8) or, if preferred, from the

TABLE 10

OBSERVATIONS AND TOTALS FOR DESIGN SR21

V_{11} 15.5	V_{13} 15.0	V_{12} 16.0	B_1 46.5	V_{22} 21.5	V_{31} 22.5	V_{43} 16.5	B_9 60.5
V_{21} 11.5	V_{22} 17.0	V_{23} 13.5	B_2 42.0	V_{12} 12.5	V_{33} 16.0	V_{41} 12.0	B_{10} 40.5
V_{33} 15.0	V_{32} 12.0	V_{31} 16.5	B_3 43.5	V_{11} 13.0	V_{23} 13.0	V_{42} 13.5	B_{11} 39.5
V_{42} 13.0	V_{43} 12.0	V_{41} 10.0	B_4 35.0	V_{13} 11.0	V_{21} 12.5	V_{32} 11.0	B_{12} 34.5
V_{21} 22.5	V_{33} 19.5	V_{42} 17.5	B_5 59.5	V_{23} 16.5	V_{32} 15.0	V_{41} 14.5	B_{13} 46.0
V_{11} 14.0	V_{32} 15.0	V_{43} 13.0	B_6 42.0	V_{13} 13.5	V_{31} 19.0	V_{42} 12.5	B_{14} 45.0
V_{13} 12.5	V_{22} 15.0	V_{41} 11.5	B_7 39.0	V_{12} 10.0	V_{21} 15.0	V_{43} 10.0	B_{15} 35.0
V_{12} 10.0	V_{23} 11.5	V_{31} 15.0	B_8 36.5	V_{11} 10.5	V_{22} 12.5	V_{33} 12.5	B_{16} 35.5
						TOTAL	G 680.5

special form in the footnote 5. To illustrate, using the latter form and values from Table 11, we show

$$t_{ij} = \frac{11}{32} T_{ij} - \frac{1}{32} \sum_{\substack{h=1 \\ h \neq j}}^{4} T_{ih} - \frac{1}{8} B_{ij.} + \frac{1}{96} G$$

and

$$t_{11} = \frac{11}{32}(53.0) - \frac{1}{32}(182.5) - \frac{1}{8}(163.5) + \frac{1}{96}(680.5) = -0.833.$$

The adjusted treatment sum of squares may be obtained through substitution in (3) or the special form in footnote 5, the unadjusted block sum of squares follows from (4), and the total sum of squares is obtained in the usual way from (5). The intra-block error is obtained by subtraction as indicated in Table 1. We show the basic analysis of variance for this example in Table 13.

Blocks in Design *SR*21 may be grouped into replications so that the design becomes a near balance, rectangular lattice. Then the unadjusted block sum of squares may be partitioned into a sum of squares for replications and a sum of squares for blocks within replications. When the blocks are grouped into rows and columns with a complete replication in each such row and column, we have a Latinized,

TABLE 11

VALUES OF T_{ij}, $B_{ij.}$, $\sum_i T_{ij}$, $\sum_j B_{ij.}$, t_{ij}, $t_{i.}$, $t_{.j}$, $\bar{t}_{i.}$ AND $\bar{t}_{.j}$ FOR THE SECOND EXAMPLE, DESIGN SR21

Values	T_{ij}				$\sum_{i=1}^{4} T_{ij}$	$B_{ij.}$				$\sum_{j=1}^{4} B_{ij.}$
	Values of j					Values of j				
i	1	2	3	4		1	2	3	4	
1	53.0	61.5	73.0	48.0	235.5	163.5	171.0	185.5	160.5	680.5
2	48.5	66.0	53.0	56.5	224.0	158.5	177.0	166.0	179.0	680.5
3	52.0	54.5	63.0	51.5	221.0	165.0	164.0	179.0	172.5	680.5

Values	t_{ij}				$t_{i.}$	$\bar{t}_{i.}$
	Values of j					
i	1	2	3	4		
1	−0.833	1.417	3.917	−2.333	2.168	0.542
2	−1.536	2.714	−0.786	−1.099	−0.707	−0.177
3	−0.943	0.120	1.432	−2.068	−1.459	−0.365
Values of $t_{.j}$	−3.312	4.251	4.563	−5.500		
Values of $\bar{t}_{.j}$	−1.104	1.417	1.521	−1.833		

rectangular lattice and the block sum of squares may be partitioned into sums of squares for rows, columns, and row by column interaction. These sums of squares, partitioning the unadjusted block sum of squares, are shown in Table 13 and the groupings of the blocks for the two special designs are indicated in Table 12. Subdivision of the unadjusted block sum of squares does not, of course, affect the rest of the intra-block analysis of variance.

TABLE 12

BLOCK ARRANGEMENTS FOR DESIGN SR21 AS A NEAR BALANCE, RECTANGULAR LATTICE AND AS A LATINIZED, RECTANGULAR LATTICE

(Entries in the table are block numbers as given in Table 10)

Near Balance, Rectangular Lattice

Replicates			
I	II	III	IV
5	15	12	3
8	13	9	1
7	16	11	4
6	14	10	2

Latinized, Rectangular Lattice

Rows	Columns			
	I	II	III	IV
I	14	2	10	6
II	16	4	12	8
III	13	1	9	5
IV	15	3	11	7

(*ii*) *Analysis for* 3 *by* 4 *Factorial.*

We may suppose that the twelve treatments of Design SR21 are made up of combinations of two factors, A and C, such that $V_{ij} = A_i C_j$, $i = 1, \cdots, 3$, $j = 1, \cdots, 4$, with A_i representing the i-th level of the A-factor and C_j, the j-th level of the C-factor. The analysis of variance is straight-forward, and may be effected through use of (22), (23), and (24) or the special forms in footnote 6. Values of $\bar{l}_{i.}$ and $\bar{l}_{.j}$ are given in Table 11 and only substitution is required. As before the adjusted AC-interaction sum of squares may be obtained by subtraction (adjusted treatment sum of squares minus the total of adjusted A-factor and C-factor sums of squares) or by direct calculation based on (24).

The analysis of variance for the 3 by 4 factorial is included in Table 15. We shall also further subdivide the factorial effects in the next subsection. We conclude this subsection with the variance estimates,

$$V(a_i - a_{i'}) = \frac{(2)(3)}{(4)(1)(12)}(1.34) = 0.17, \qquad i \neq i',$$

TABLE 13

Basic Analyses of Variance, Design SR21

Source	D.f.	S.s.	M.s.	F
Treatments (adj.)	11	115.15	10.47	7.81
Blocks (unadj.)	15	311.91	20.79	....
Subdivision for Near Balance, Rectangular Lattice				
Replicates	3	12.94	4.31	....
Blocks in Replicates	12	298.98	24.92	....
Subdivision for Latinized, Rectangular Lattice				
Replicates (columns)	3	12.94	4.31	....
Rows	3	232.31	77.44	....
Rows by Columns	9	66.67	7.41	....
Intra-Block Error	21	28.19	1.34	....
Total	47	455.25	....	....

and

$$V(c_j - c_{j'}) = \frac{(2)(3)}{3(0 + 12 - 4)}(1.34) = 0.34, \qquad j \neq j',$$

obtainable from (25) and (26). These variances respectively apply to all contrasts among A-factor and C-factor effects and on this basis are appropriate for properly selected contrasts of the multi-factor factorial of the next subsection.

(iii) Analysis for 3 by 2^2 Factorial.

To complete the illustration of the use of factorials in a near balance design, we now consider the treatments to be the treatment combinations of a 3 by 2^2 factorial and subdivide the A-factor, C-factor, and AC-interaction sums of squares accordingly. We take the C-factor to be divisible into four combinations of two-level factors N and P. The identification of the treatments V_{ij} in terms of these new factors is shown in Table 14. In that table we also give the coefficients for

TABLE 14

COEFFICIENTS FOR THE ANALYSIS OF DESIGN SR21 AS A 3 BY 2^2 FACTORIAL

Treatments	V_{11}	V_{12}	V_{13}	V_{14}	V_{21}	V_{22}	V_{23}	V_{24}	V_{31}	V_{32}	V_{33}	V_{34}	
3 by 2^2 Factorial designations	$A_1N_1P_1$	$A_1N_2P_1$	$A_1N_1P_2$	$A_1N_2P_2$	$A_2N_1P_1$	$A_2N_2P_1$	$A_2N_1P_2$	$A_2N_2P_2$	$A_3N_1P_1$	$A_3N_2P_1$	$A_3N_1P_2$	$A_3N_2P_2$	Sums of squared coefficients
Treatment estimators	t_{11} = −0.833	t_{12} = 1.417	t_{13} = 3.917	t_{14} = −2.333	t_{21} = −1.536	t_{22} = 2.714	t_{23} = −0.786	t_{24} = −1.099	t_{31} = −0.943	t_{32} = 0.120	t_{33} = 1.432	t_{34} = −2.068	
Contrasts													
Linear A	−1	−1	−1	−1	0	0	0	0	+1	+1	+1	+1	8
Quadratic A	−1	−1	−1	−1	+2	+2	+2	+2	−1	−1	−1	−1	24
N	−1	+1	−1	+1	−1	+1	−1	+1	−1	+1	−1	+1	12
P	−1	−1	+1	+1	−1	−1	+1	+1	−1	−1	+1	+1	12
NP	+1	−1	−1	+1	+1	−1	−1	+1	+1	−1	−1	+1	12
$N \times$ Lin. A	+1	−1	+1	−1	0	0	0	0	−1	+1	−1	+1	8
$N \times$Quad.A	+1	−1	+1	−1	−2	+2	−2	+2	+1	−1	+1	−1	24
$P \times$ Lin. A	+1	+1	−1	−1	0	0	0	0	−1	−1	+1	+1	8
$P \times$Quad.A	+1	+1	−1	−1	−2	−2	+2	+2	+1	+1	−1	−1	24
$NP \times$Lin.A	−1	+1	+1	−1	0	0	0	0	+1	−1	−1	+1	8
$NP \times$Quad. A	−1	+1	+1	−1	+2	−2	−2	+2	−1	+1	+1	−1	24

the indicated contrasts and the sums of these squared coefficients for each row. The sums of squares for the linear contrasts are computed as in complete block analysis except that additional multipliers are required as shown in (32), (34), and (35). The sums of squares for the contrasts used are given in Table 15. We have illustrated this type of computation in the first example and do not do so again here.

Table 15 is sufficient to complete our illustration of the use of factorials in this second example. The unadjusted block sum of squares is not subdivided in Table 15 but the subdivisions of Table 13 may be used when desired.

TABLE 15

ANALYSIS OF VARIANCE FOR THE 3 BY 2^2 FACTORIAL

Source	D.f.	S.s.	M.s.	F
Treatments (adj.)	11	115.15	10.47	7.81
A-Factor	2	7.32	3.66	2.73
Linear A	1	6.57	6.57	4.90
Quad. A	1	0.75	0.75	....
C-Factor	3	71.20	23.73	17.7
N	1	1.39	1.39	1.04
P	1	0.78	0.78	....
NP	1	69.03	69.03	51.51
AC-Interaction	6	36.63	6.10	4.55
$N \times$ Lin. A	1	0.81	0.81	....
$N \times$ Quad. A	1	22.76	22.76	16.99
$P \times$ Lin. A	1	0.22	0.22	....
$P \times$ Quad. A	1	5.94	5.94	4.43
$NP \times$ Lin. A	1	5.17	5.17	3.86
$NP \times$ Quad. A	1	1.72	1.72	1.28
Blocks (unadj.)	15	311.91	....	....
Intra-Block Error	21	28.19	1.34	....
Total	47	455.25	....	....

6. DISCUSSION AND SUMMARY

Incomplete block designs have been widely used in agronomic and animal experimentation but only to a limited extent, until recently, in industrial and engineering research. In the latter areas it may be that the frequent need to use factorial arrangements of treatments has in some measure precluded the use of incomplete block designs. In

this paper, we have shown how factorials may be used in group divisible, partially balanced, incomplete block designs and illustrated the appropriate, and quite simple, method of analysis.

Factorials in incomplete block designs should be widely useful. In agriculture, they appear to be particularly useful in animal experimentation, where it is only possible to obtain small groups of homogeneous animals, as from a litter, and these may constitute material for an incomplete block. In large animal experimentation, twins may be used in some cases in designs with blocks of size two. In experiments involving subjective judgments, it is necessary to keep the block sizes small owing to fatigue and adaption effects. There are many situations where factorials in incomplete block designs will be useful in taste testing; the only limitation is that our methods require applications where scores, that may reasonably be used with an assumption of normality, are obtainable. In industrial research, it frequently happens that there are limitations on the number of samples or runs that may be obtained as a homogeneous group. Limiting factors are time, batch sizes, equipment, sources of raw material, and personnel to cite a few, and incomplete block designs are suggested to circumvent such limitations.

In this paper we have shown how m by n factorials may be introduced into group divisible designs with m classes of n items. The introduction of factorial treatments does not affect the fact that some of these designs, such as near balance, and Latinized, rectangular lattices, may incorporate grouping of the blocks into replications or into a Latin square. Harshbarger [4], in discussing rectangular lattices, suggested that the $K(K - 1)$ series were of primary importance in that they provided treatment numbers well positioned between those available from square lattice designs. All series with $K(K - Q)$ treatments become important in making factorials, with a wide variety of factor levels, available. It should be further noted that, while Q is an integer, it may in fact be either a positive or negative integer and designs with Q negative are included in the cited catalogue of designs.

The adjusted treatment sum of squares for the group divisible designs was obtained as a simple function of the treatment-effect estimators in the development of the method for factorials. This, in itself, seems to be a useful new result and one that is helpful to an understanding of the analysis of such designs. The analysis for factorials was effected in terms of this form of the adjusted treatment sum of squares and in fact depends on a simple partitioning of that quantity. The main theoretical discussions are for an m by n factorial, but it is further demonstrated how multi-factor factorials may be used when

the factor levels are divisors of m or n. Single degree of freedom comparisons have been made in much the usual way. Illustrations in the examples were chosen to show a variety of meaningful multi-factor factorials along with single degree of freedom comparisons.

We did not consider how group divisible designs may be generated. Instead reference is made to a catalogue of designs and references there lead to methods of generating such designs.

Recovery of inter-block information is being considered by R. E. Walpole working with the present authors. While this additional work is not yet available, considerable progress has been made. Research is also well under way on the problem of incorporating factorials into other partially balanced, incomplete block designs. Some of these other classes of designs are given by Bose, Clatworthy, and Shrikhande [3] and are the Simple, Triangular, and Latin Square types of two-associate class designs. Extensions to designs with more than two associate classes appear to be straight-forward and these designs may yield more flexibility for use with factorials.

We have illustrated the method of analysis for factorials for two designs with numerical data. In doing this, we have carried out the basic analysis of variance in each case and then subdivided the adjusted treatment sum of squares into sums of squares for factorial effects. We have concluded the examples with the analysis of variance tables. It should be noted that, in analyses and reports on actual experimental data, care should be taken, as usual, to further summarize and interpret the findings of the experiment. Discussion of an experiment should not terminate with an analysis of variance table; we have not included interpretation of results because we were primarily interested in presenting new techniques of analysis, and interpretation follows in the usual way for factorial experiments.

We are pleased to acknowledge the assistance of Boyd Harshbarger who offered helpful suggestions incorporated in this paper and closely followed the progress of this research.

REFERENCES

[1] Kramer, C. Y., *Factorial Treatments in Incomplete Block Designs*, Ph.D. Dissertation, Virginia Polytechnic Institute Library, Blacksburg, Va., 1956.

[2] Kramer, C. Y. and Bradley, R. A., Intra-Block Analysis for Factorials in Two-Associate Class, Group Divisible Designs, *Annals Math. Stat., 28*, No. 2, 1957.

[3] Bose, R. C., Clatworthy, W. H., and Shrikhande, S. S., *Tables of Partially Balanced Incomplete Block Designs with Two Associate Classes*, Tech. Bul. No. 107, N. C. Agr. Exp. Stat., 1954.

[4] Clatworthy, W. H., Partially Balanced Incomplete Block Designs with Two Associate Classes and Two Treatments Per Block, *J. of Research of Nat. Bur. of Standards, 54*, 177–190, 1955.

[5] Harshbarger, Boyd, Near Balance Rectangular Lattices, *Va. J. of Sci. 2* (new series), 13–27, 1951.

[6] Harshbarger, Boyd, Latinized Rectangular Lattices, *Biometrics 8*, 73–84, 1952.

[7] Harshbarger, Boyd, *The 2^3 Factorial in a Latinized Rectangular Lattice Design*, Tech. Rept. 1 to Office of Ordnance Research, OOR Project No. 1166, 1954.

[8] Bose, R. C. and Nair, K. R., Partially Balanced Incomplete Block Designs, *Sankhyā, 4*, 337–372, 1939.

[9] Bose, R. C. and Shimamoto, T., Classification and Analysis of Partially Balanced Incomplete Block Designs with Two Associate Classes, *J. of Amer. Stat. Assn., 47*, 151–184, 1952.

[10] Cornish, E. A., Factorial Treatments in Incomplete Randomized Blocks, *J. Australian Inst. Agr. Sci., 4*, 199–203, 1938.

REPEATED LINEAR REGRESSION AND VARIANCE COMPONENTS OF A POPULATION WITH BINOMIAL FREQUENCIES

C. C. Li

Department of Biostatistics, University of Pittsburgh, Pittsburgh, Pennsylvania, U. S. A.

1. *The Initial Problem*

Consider the variate $X = 0, 1, \cdots, i, \cdots, n$, whose probability density function is

$$\text{Prob}\,(X = i) = f_i = \frac{n!}{i!(n-i)!}\, p^i q^{n-i}.$$

It is well known that $\bar{X} = np$ and $\sigma_X^2 = npq$. Now suppose that there is a concomitant variate Y that takes the value Y_0 when $X = 0$, takes the value Y_1 when $X = 1$, and so on. Therefore, f_i is also the frequency of the paired values $(X, Y) = (i, Y_i)$. Our problem is to find the linear regression coefficient of Y on X. The least square method leads to the solution that the slope of the fitted straight line should be

$$b = \sigma_{XY}/\sigma_X^2\,,$$

where σ_{XY} denotes the covariance of X and Y. Since it is already known that $\sigma_X^2 = npq$, our major task then is to find the value of $\sigma_{XY} = \sum fXY - \bar{X}\bar{Y}$.

2. *The Theorem*

Recall that the frequencies of $Y_0, Y_1, \cdots, Y_n$ are given by the terms f_i of $(q + p)^n$. Now denote the successive differences

$$(Y_1 - Y_0), \quad (Y_2 - Y_1), \cdots, (Y_n - Y_{n-1})$$

by

$$h_0\,, \qquad h_1\,, \qquad \cdots\,, \qquad h_{n-1}$$

as illustrated in Figure 1. These h's are the slopes of the segments joining two adjacent points. The theorem, to be proved after consider-

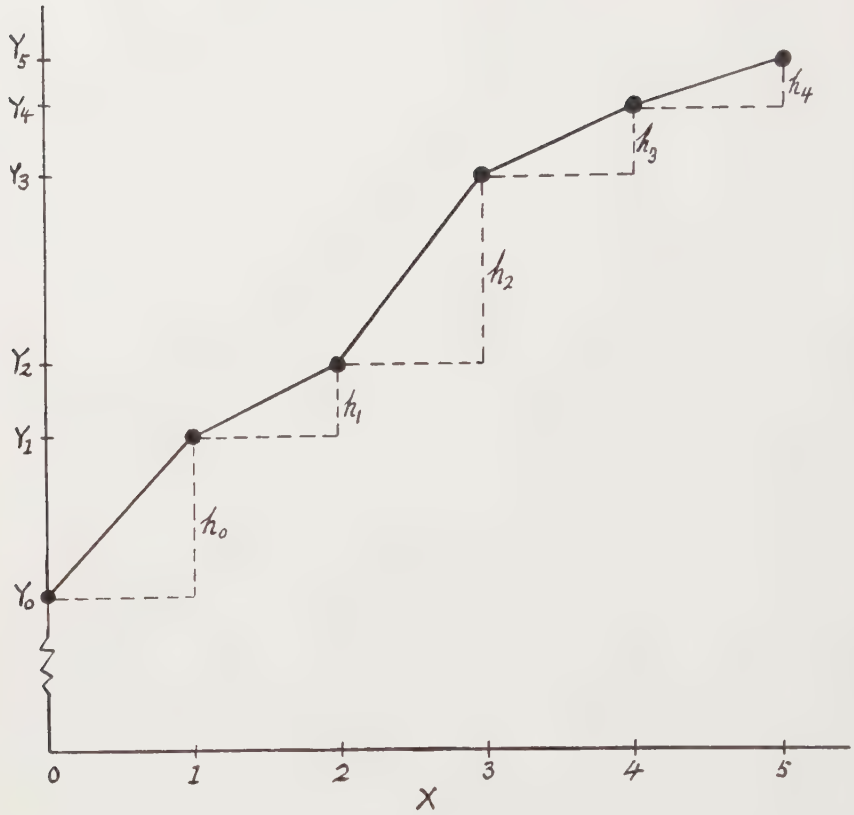


FIGURE 1

THE REGRESSION OF Y ON X WHEN THE FREQUENCIES OF THE PAIRED VALUES OR POINTS ARE GIVEN BY EXPANDING $(q + p)^n$. THE DIFFERENCE $h_i = Y_{i+1} - Y_i$ IS THE SLOPE OF THE SEGMENT JOINING TWO ADJACENT POINTS AS X VARIES BY UNIT STEPS. THE SLOPE OF THE LEAST SQUARE LINE (NOT SHOWN IN THE GRAPH) IS THE WEIGHTED MEAN OF THE h'S WHERE THE WEIGHTS ARE TERMS OF $(q + p)^{n-1}$.

ing an example, asserts that the regression coefficient of Y on X is simply the weighted mean of the segment slopes, where the weights, f_i' , are terms of $(q + p)^{n-1}$; that is:

$$
\begin{aligned}
b &= \sum_{i=0}^{n-1} \binom{n-1}{i} p^i q^{n-1-i} (Y_{i+1} - Y_i) \\
&= q^{n-1}h_0 + (n-1)pq^{n-2}h_1 + \cdots + p^{n-1}h_{n-1} \\
&= \sum f_i' h_i \, .
\end{aligned}
$$

3. *An Example*

Let us take the example in which $n = 5$. Then the values of f_i are the six terms of $(q + p)^5$ and $\bar{X} = 5p$ and $\sigma_X^2 = 5pq$. The purpose of giving such an example at this stage is not so much to show the end result as to outline the method of proof employed in the next section. By following the steps of Table 1, much explanation may be saved and possible confusion may be avoided later. The first column of Table 1 gives

$$\sum fXY = 5p(q^4Y_1 + 4pq^3Y_2 + \cdots + p^4Y_5) = 5p\bar{Y}',$$

where $\bar{Y}'$ is the weighted mean of Y_1, Y_2, $\cdots$, Y_5, with the terms f_i' of $(q + p)^4$ as weights. Thus the covariance is

$$\sigma_{XY} = \sum fXY - \bar{X}\bar{Y} = 5p(\bar{Y}' - \bar{Y})$$

We can now ignore the factor $\bar{X} = 5p$ and concentrate on the part $\bar{Y}' - \bar{Y}$. To facilitate simplification, we split each term of $\bar{Y}'$ into two parts by multiplying it by $(p + q) = 1$. Take, for instance, the term involving Y_3:

$$6p^2q^2Y_3(p + q) = 6p^3q^2Y_3 + 6p^2q^3Y_3 .$$

Then, as indicated in the middle column of Table 1, put the first term on the same line of the original term but put the second term in the

TABLE 1

CALCULATION OF THE COVARIANCE IN A $(q + p)^5$ POPULATION.

Product term fXY	Ignore $5p$; split by $(p + q)$	Minus terms of $\bar{Y}$
$q^5 0Y_0 = 0$	q^5Y_1	$-q^5Y_0$
	+	
$5pq^4 1Y_1 = 5p \cdot q^4Y_1$	pq^4Y_1 $4pq^4Y_2$	$-5pq^4Y_1$
	+	
$10p^2q^3 2Y_2 = 5p \cdot 4pq^3Y_2$	$4p^2q^3Y_2$ $6p^2q^3Y_3$	$-10p^2q^3Y_2$
	+	
$10p^3q^2 3Y_3 = 5p \cdot 6p^2q^2Y_3$	$6p^3q^2Y_3$ $4p^3q^2Y_4$	$-10p^3q^2Y_3$
	+	
$5p^4q 4Y_4 = 5p \cdot 4p^3qY_4$	$4p^4qY_4$ p^4qY_5	$-5p^4qY_4$
	+	
$p^5 5Y_5 = 5p \cdot p^4Y_5$	p^5Y_5	$-p^5Y_5$
Total: $5p \cdot \bar{Y}'$	$\bar{Y}'$	$-\bar{Y}$

preceding line involving Y_2 . Finally, write down the terms of $\bar{Y}$ in the usual order in the last column. Then the value of $(\bar{Y}' - \bar{Y})$ is the sum of the differences obtained by subtracting the corresponding terms (on the same line) of the last column from those of the middle column. Thus, we obtain

$$\bar{Y}' - \bar{Y} = q^5(Y_1 - Y_0) + 4pq^4(Y_2 - Y_1) + \cdots + p^4q(Y_5 - Y_4)$$
$$= q\{q^4h_0 + 4pq^3h_1 + 6p^2q^2h_2 + 4p^3qh_3 + p^4h_4\}.$$

Consequently,

$$\sigma_{XY} = 5p(\bar{Y}' - \bar{Y}) = 5pq\{q^4h_0 + \cdots + p^4h_4\}$$

and

$$b = \sigma_{XY}/5pq = q^4h_0 + 4pq^3h_1 + \cdots + p^4h_4$$

as asserted by the theorem previously.

4. *General Demonstration*

The following demonstration for the general truth of the theorem follows precisely the steps outlined in the preceding example. It is merely a rewriting of the relationship in a more general form. Maybe it is not the best proof that can be given, but it requires only the knowledge of a well known identity of binomial coefficients. The covariance for any integral value of n is

$$\sigma_{XY} = \sum fXY - \bar{X}\bar{Y}$$
$$= \sum \frac{n!}{i!(n-i)!} p^i q^{n-i} i Y_i - np\bar{Y}$$
$$= np\left\{\sum \frac{(n-1)!}{(i-1)!(n-i)!} p^{i-1} q^{n-i} Y_i - \bar{Y}\right\}$$

Split each term of the summation into two by multiplying it by $(p + q)$; thus the two successive terms involving Y_i and Y_{i+1} become, respectively:

$$\frac{(n-1)!}{(i-1)!(n-i)!} p^i q^{n-i} Y_i + \frac{(n-1)!}{(i-1)!(n-i)!} p^{i-1} q^{n-i+1} Y_i$$

and

$$\frac{(n-1)!}{i!(n-i-1)!} p^{i+1} q^{n-i-1} Y_{i+1} + \frac{(n-1)!}{i!(n-i-1)!} p^i q^{n-i} Y_{i+1}$$

The corresponding term in $\bar{Y}$ (which is to be subtracted) is

$$-\frac{n!}{i!(n-i)!}p^i q^{n-i}Y_i .$$

Collecting the three terms involving the same $p^i q^{n-i}$ and noting that

$$\frac{n!}{i!(n-i)!} - \frac{(n-1)!}{(i-1)!(n-i)!} = \frac{(n-1)!}{i!(n-i-1)!} ,$$

we obtain the term

$$\frac{(n-1)!}{i!(n-i-1)!}p^i q^{n-i}(Y_{i+1} - Y_i) = q\binom{n-1}{i}p^i q^{n-i-1}h_i .$$

Summing such terms, we obtain the covariance

$$\sigma_{XY} = npq \sum f'_i h_i$$

where f'_i are the terms of $(q+p)^{n-1}$. Dividing σ_{XY} by $\sigma_X^2 = npq$, we establish the theorem that the linear regression coefficient of Y on X is the weighted mean of the segment slopes h_i where the weights are terms of $(q+p)^{n+1}$ as stated in section 2.

Having found the slope of the regression line, we can write down the equation of the straight line. Furthermore, we know that the portion of σ_Y^2 that is due to regression is $b\sigma_{XY} = b^2\sigma_X^2$, and in this case it is equal to $npqb^2$.

5. *Repeated linear regression*

In order to continue the process of linear regression in successive stages, we need a more general system of notation. If Y is the original series of values, we shall use Y' to denote the successive differences of Y so that $Y'_0 = Y_1 - Y_0$, etc., which has been denoted by h in our preliminary theorem. Similarly, $Y''_0 = Y'_1 - Y'_0$, etc. The procedure is outlined in Table 2 for the case $n = 5$. An analogous system of notation may be adopted for the frequencies. Thus, f denotes the terms of $(q+p)^n$, and f' the terms of $(q+p)^{n-1}$, etc. Further, the mean of the original Y values is $\bar{Y} = \sum fY$, and the mean of the Y' values is $\bar{Y}' = \sum f'Y'$, etc. as indicated at the bottom of Table 2. What we have shown in the preceding section is that the linear regression coefficient in the original population (*i.e.* the (0)-column) is equal to the mean of the next column to the right (*i.e.* the (1)-column). From the theorem we have demonstrated, it follows that the regression coefficient for the (1)-column is equal to the mean of the (2)-column, etc. Hence, the successive mean values $\bar{Y}'$, $\bar{Y}''$, $\cdots$, are all regression coefficients.

TABLE 2

REPEATED REGRESSION FOR BINOMIALS

(0)		(1)		(2)		(3)		(4)		(5)	
Y	f	Y'	f'	Y''	f''	Y'''	f'''	Y''''	f''''	Y^{V}	f^{V}
Y_0	q^5										
		Y'_0	q^4								
Y_1	$5pq^4$			Y''_0	q^3						
		Y'_1	$4pq^3$			Y'''_0	q^2				
Y_2	$10p^2q^3$			Y''_1	$3pq^2$			Y''''_0	q		
		Y'_2	$6p^2q^2$			Y'''_1	$2pq$			Y^{V}_0	1
Y_3	$10p^3q^2$			Y''_2	$3p^2q$			Y''''_1	p		
		Y'_3	$4p^3q$			Y'''_2	p^2				
Y_4	$5p^4q$			Y''_3	p^3						
		Y'_4	p^4								
Y_5	p^5										
$\bar{Y}$		$\bar{Y}'$		$\bar{Y}''$		$\bar{Y}'''$		$\bar{Y}''''$		$\bar{Y}^{\mathrm{V}}$	

6. *Components of Variance of Y*

We have already noted that one portion of the variance of Y is $b\sigma_{XY} = b^2\sigma_X^2 = npqb^2 = npq\bar{Y}'^2$. Now for the case of $n = 5$, for example, we note that the following expression is an identity:

$$\sum fY^2 = \bar{Y}^2 + 5pq\bar{Y}'^2 + 10p^2q^2\bar{Y}''^2 + 10p^3q^3\bar{Y}'''^2 + 5p^4q^4\bar{Y}''''^2 + p^5q^5\bar{Y}^{\mathrm{V}^2}$$

For the general case of the $(q + p)^n$ population we have

$$\sum fY^2 = \bar{Y}^2 + npq\bar{Y}'^2 + \binom{n}{2}p^2q^2\bar{Y}''^2 + \binom{n}{3}p^3q^3\bar{Y}''''^2 + \cdots$$

In other words, the variance $\sigma_Y^2 = \sum fY^2 - \bar{Y}^2$ may be partitioned into n components. The first component is due to variation in the original Y values as expressed in terms of $Y'_i = Y_{i+1} - Y_i$ where $i = 0, 1, 2, \cdots$. The second component is due to the variation of Y' expressed in terms $Y''_i = Y'_{i+1} - Y'_i$; or the variation of the original Y in terms of $Y_{i+2} - 2Y_{i+1} + Y_i$. Similarly, the third component is due to $Y_{i+3} - 3Y_{i+2} + 3Y_{i+1} - Y_i$, and so on. In view of this property, we may speak of the first component as the linear component, the second quadratic, the third cubic, etc.

7. *Outline of Proof*

For small values of n, the above identity may be directly verified by writing out every term of $\bar{Y}^2$, $\bar{Y}'^2$, $\bar{Y}''^2$, $\cdots$ and seeing that both sides are equal. For the general case we need to show that the coefficient of Y_i^2, after collecting the terms of the right hand side, reduces to $f_i = \binom{n}{i} p^i q^{n-i}$ and that the coefficient of $Y_i Y_j$ is zero ($i \neq j$). This will involve a great deal of writing. Presumably the identity may be established through more advanced mathematical tools which the author does not possess. The following indicates the method of a longhand proof and suffices to show that the coefficient of Y_0^2 is q^n and that of $Y_0 Y_1$ is zero.

Source	Coefficient of Y_0^2	Coefficient of $Y_0 Y_1$
$\bar{Y}^2$	q^{2n}	$2npq^{2n-1}$
$npq\bar{Y}'^2$	npq^{2n-1}	$-2npq^{2n-1} + 2n(n-1)p^2q^{2n-2}$
$\binom{n}{2} p^2q^2\bar{Y}''^2$	$\binom{n}{2} p^2q^{2n-2}$	$-2n(n-1)p^2q^{2n-2} + \cdots$
$\cdots$	$\cdots$	$\cdots$
$np^{n-1}q^{n-1}\bar{Y}^2$ (with $n-1$ primes)	$np^{n-1}q^{n+1}$	$-2n(n-1)p^{n-1}q^{n+1} + 2np^nq^n$
$p^nq^n\bar{Y}^2$ (with n primes)	p^nq^n	$-2np^nq^n$
sum	q^n	0

8. *Corollaries*

From the general identity and the example given in Table 2 it is clear that some of the components of σ_Y^2 may be zero. For instance, if the Y''' values in the (3)-column are all zero, which implies that the Y'' values are a constant, which in turn implies that the Y' values form an arithmetic progression, then σ_Y^2 has only the first two components, while all subsequent ones are zero.

This example, however, does not mean that σ_Y^2 always has the first k components and the subsequent $n - k$ components vanish. A more general statement is that if the mean of a certain column is zero, the corresponding component of the variance of Y vanishes. In particular, if $p = q = \frac{1}{2}$, and the original Y values are symmetrical, such as

$$Y\colon \quad 2, \quad 4, \quad 8, \quad 8, \quad 4, \quad 2$$

for $n = 5$, it will be found that the first, third, and fifth components vanish and the variance of Y consists of only the second and the fourth components. Finally, as an extreme example, if the original Y values constitute an arithmetic series, so that Y' is a constant, then $\sigma_Y^2 = npq\bar{Y}'^2$, having only one component. Further, if $Y' = 1$, $\sigma_Y^2 = npq$, which is the ordinary binomial variance.

9. *An Application in Genetics*

In a random mating autotetraploid population with random chromosome segregation, the proportions of the five genotypes with respect to one pair of genes are as follows:

Genotype :	$aaaa$	$Aaaa$	$AAaa$	$AAAa$	$AAAA$
Frequencies :	q^4	$4pq^3$	$6p^2q^2$	$4p^3q$	p^4
Measurement:	Y_0	Y_1	Y_2	Y_3	Y_4

The Y values are the measurements of a quantitative trait (such as weight, vitamin content, oil content, etc.) of the various genotypes. If the trait under consideration is subject to random fluctuations, we take the mean value of the trait for each genotype as the "genotypic value" of that genotype and treat it as our Y. The variance of Y in a population is known as the "genotypic variance".

Applying our theorem on the components of variance of binomial populations and putting $n = 4$, we may split the genotypic variance into four components (Li [1957]). The first component is due to differences of the type $Y_1 - Y_0$ or $Aaaa - aaaa$. The second component is due to the values of the type $Y_2 - 2Y_1 + Y_0$, or $AAaa - 2Aaaa + aaaa$. The third component is due to the discrepancies of the type $Y_3 - 3Y_2 + 3Y_1 - Y_0$, or $AAAa - 3AAaa + 3Aaaa - aaaa$. The last component is due to $AAAA - 4AAAa + 6AAaa - 4Aaaa - aaaa$. These components have been named the additive, digene, trigene, and quadrigene components (Kempthorne [1954]). A numerical example is given in Table 3. The partition of the genotypic variance into such components will facilitate the prediction of the correlation between relatives, especially the parent-offspring and the sib-sib correlations which are useful in breeding work. A more detailed discussion of the subject would be out of place here.

REFERENCES

Kempthorne, O. [1954]. The correlation between relatives in a simple autotetraploid population. *Genetics 40*: 168–174.

Li, C. C. [1957]. The genetic variance of autotetraploids with two alleles. *Genetics 42*: In press.

TABLE 3

GENOTYPIC VARIANCE COMPONENTS OF AN AUTOTETRAPLOID F_2 POPULATION WITH TWO ALLELES ($p = q = \frac{1}{2}$).

Y = Average Measurement of Genotype.

Genotype	Y	f	Y'	f'	Y''	f''	Y'''	f'''	Y''''	f''''
aaaa	8	$\frac{1}{16}$								
			12	$\frac{1}{8}$						
Aaaa	20	$\frac{4}{16}$			16	$\frac{1}{4}$				
			28	$\frac{3}{8}$			-28	$\frac{1}{2}$		
AAaa	48	$\frac{6}{16}$			-12	$\frac{2}{4}$			16	1
			16	$\frac{3}{8}$			-12	$\frac{1}{2}$		
AAAa	64	$\frac{4}{16}$			-24	$\frac{1}{4}$				
			-8	$\frac{1}{8}$						
AAAA	56	$\frac{1}{16}$								
mean	$\bar{Y} = 43$		$\bar{Y}' = 17$		$\bar{Y}'' = -8$		$\bar{Y}''' = -20$		$\bar{Y}'''' = 16$	

additive component	$= 4pq\bar{Y}'^2 = 4(\frac{1}{2})(\frac{1}{2})(17)^2$	$= 289$
digene component	$= 6p^2q^2\bar{Y}''^2 = 6(\frac{1}{2})^2(\frac{1}{2})^2(-8)^2$	$= 24$
trigene component	$= 4p^3q^3\bar{Y}'''^2 = 4(\frac{1}{2})^3(\frac{1}{2})^3(-20)^2$	$= 25$
quadrigene component	$= p^4q^4\bar{Y}''''^2 = (\frac{1}{2})^4(\frac{1}{2})^4(16)^2$	$= 1$
Total genotypic variance	$= \sigma_Y^2 =$ sum	$= 339$

APPRECIATION

In January 1957, with the final transfer of Treasurer's duties, Dr. Chester I. Bliss terminated over nine years of sterling service as a general officer of our International society.

I doubt very much whether the membership fully appreciates the extent of its debt to Dr. Bliss. The Biometric Society was organised in September 1947 at the Marine Biological Laboratory in Woods Hole, Massachusetts, during the First International Biometric Conference. This conference was arranged on the initiative of the Biometrics Section of the American Statistical Association, and one of its objectives was to set international cooperation in biometry on an effective and enduring foundation. Many individuals were inspired by this objective, but it was Chester Bliss who, besides the inspiration, had the enthusiasm and driving energy so necessary to bring the plan to fruition. Bliss played the major rôle in organising the Woods Hole meeting, and I believe it is fair to say that he, more than any other, was responsible for the foundation of the Society. His convictions were amply justified. Within two years, 6 organised regions were established, and the membership totalled nearly 900. Four years later, 2 more regions had been added, and the membership had been increased by a further 25 per cent; today, there are 9 organised regions, and 6 other national groups, covering over 1300 members.

Such rapid expansion made huge demands on Bliss, first as Secretary and later as Secretary-Treasurer, but he still found time to play his part in the organisation of three international biometric conferences and two international biometric symposia. This record of unselfish devotion to the Society speaks for itself.

I count it a great privilege to express, on behalf of the membership, our deep appreciation and grateful thanks to Chester Bliss.

E. A. Cornish
President

QUERIES AND NOTES

GEORGE W. SNEDECOR, *Editor*

Arrangements of Pots in Greenhouse Experiments

126 QUERY: The proper method of arranging pots in a greenhouse pot experiment has always been a source of worry and confusion to me. I am speaking specifically about greenhouse tests in which the pots are arranged in a randomized block design in order to secure yield, or other measurement data to be subjected to an analysis of variance.

Is it proper, or is it improper, to periodically rearrange the pots in a randomized fashion within a replication? If the pots were field plots, it would, of course, be impossible to periodically rearrange them. Yet, it has been my observation that many agronomists periodically rearrange the location of the pots within the replications. The argument of agronomists is that periodical rearrangement will reduce border, heat and light effects. Perhaps the rearrangements only spread these environmental effects throughout the entire replication. If this is true, is it desirable and does it lend to the accuracy of the experiment?

Another question concerns the rearranging of the location of the entire replication. Are we justified in periodically exchanging the locations of the entire replications? I would certainly be happy to have any information or opinions on this problem.

ANSWER: The query is partially answered by the following quotation [Designs of greenhouse experiments for statistical analysis, G. Cox and W. G. Cochran, *Soil Science, 62,* 87–98, 1946]:

"In the foregoing experiments, the pots within the incomplete blocks were re-randomized at intervals during the early part of the trial. It is now believed that this was unnecessary. ··· Unfortunately, we have found no data on the effectiveness of the practice: its drawbacks are the labor involved, the possibility of injury to the plants, and the opportunity for unobserved biases. The use of incomplete block designs should render the practice of less value ··· ."

In addition I would make the following comments. First, however, I would say that if there is any possibility of injury to the plants the pots

should not be moved around. With this qualification it seems to me that the two relevant aspects are (1) the validity of the experiment with relocating of pots and (2) efficiency of so doing.

My opinion is that the validity of the experiment is unaffected by moving the pots around in either a prechosen way independent of the treatments or at random. We may regard the total error of a pot yield as being made up of the following parts:

(1) errors of measurements of yield
(2) errors peculiar to the pot, e.g. the soil or matrix, the plant, and possibly even a peculiarity of the pot *per se*
(3) errors due to location of the pot in the greenhouse

Parts (1) and (2) are irrelevant to or invariant in the consideration of your questions. As regards (3), we can visualize the total error as consisting of the total effect of deviations of the location of the pot from the average in the block for as many time intervals as we like. These we would expect to be non-additive in their effects on yield but we would expect that if a location gets more sunshine for example in the first month it will get more in the second month and so on. By moving the pots around I would expect that we would more nearly equalize the effects of location of pots among the treatments and hence reduce the experimental error. It might be argued that there would be some interaction with location in the block. This will, however, behave partially as random error and keeping the pots in the same location in the block will not make such an effect easier to discover.

As regards efficiency we may envisage components of error variance as follows: σ_1^2 (measurement), σ_2^2 (pot), σ_3^2 (location). The total error will be

$$\sigma_1^2 + \sigma_2^2 + \sigma_3^2$$

Perhaps by moving the pots around one halves the last component, in which case the total error would be

$$\sigma_1^2 + \sigma_2^2 + \frac{\sigma_3^2}{2}$$

You can visualize that if, for example,

$$\sigma_1^2 = 5, \qquad \sigma_2^2 = 25, \qquad \sigma_3^2 = 5$$

one will not achieve much greater efficiency by moving pots around particularly if a lot of work is involved. If on the other hand

$$\sigma_1^2 = 5, \qquad \sigma_2^2 = 25, \qquad \sigma_3^2 = 50$$

and the moving does cut the third component in half, we would reduce the error variance from 80 to 55 so that we would increase the efficiency of the experiment by

$$100\left(\frac{80}{55} - 1\right) \text{ per cent or } 45 \text{ per cent.}$$

We would need for example 4 replicates, with pots moved around, to obtain as much information as we would get with 6 replicates and no moving of pots. Of course some labor would be involved in doing the moving, which might better be spent on more replicates.

On your second question of relocating whole replicates, there are no logical difficulties at all. We may for instance visualize block 1 as being on bench 1 for the first month, on bench 3 for the second month and on bench 2 for the third month. We would tend, it seems to me, to reduce block-treatment interactions by doing this. You may recall that an assumption necessary for the analysis of randomized blocks is that there be no block-treatment interaction, unless the blocks are a sample from a very large population of blocks (which is rare in my opinion). Clearly relocating blocks will result in blocks being more nearly alike which will tend to minimize the block-treatment interactions. We may also expect that this will at the same time result in a decrease in pure error variance because there will be a positive correlation between the location-within-block errors associated with any one pot during the time intervals unless the pots are moved. My opinion of your second question is therefore that to relocate blocks will almost certainly reduce error and make the usual analysis more nearly valid but the gain in efficiency may again be very slight.

To close, one may suspect that if there were sizeable gains in net efficiency the procedure would have been used widely. On the other hand I know of no experimental work on the question.

Oscar Kempthorne
Iowa State College
Ames, Iowa, U.S.A.

ABSTRACTS

Papers presented at the joint meeting of the Biometric Society (Brazilian Region) and the Brazilian Association for the Advancement of Science, Escola de Minas e Metalurgia, Ouro Preto, Minas Gerais, July 6, 1956

402 C. G. FRAGA, JR., AND R. MEIRELLES DE MIRANDA (Instituto Agronômico, Campinas). **Analysis of a Non-Orthogonal Experiment.**

Methods for the analysis of non-orthogonal experiments are discussed and the process of fitting constants is considered in detail. This process was used in the analysis of a feeding trial with chickens. Four treatments were compared and five replicates were used, each consisting of ten birds, except one, which had nine chickens. The ratio of males to females in each replicate was variable. The constants were computed by least squares and also by using a successive approximation method due to W. L. Stevens (*Biometrika*, 1948).

403 J. SANTOS DANIEL, ITA R. K. ABRAMOF AND T. SILVA (Instituto Agronômico, Belo Horizonte). **Statistical Analysis of the Number of Stomata in Coffea Under Different Experimental Conditions.**

The variation in the number of stomata was analyzed for 3 different parts of the same leaves, different leaves of the same plant, different plants of the same variety and 20 different varieties. The effects of shadowing and irrigation were also studied. The significance of main effects and interactions was evaluated by a series of analyses of variance. Significant variation was found among varieties and between different parts of the same leaves in some varieties. The observed decrease in the number of stomata under shadowing was highly significant, and the interaction of this effect with that of variety was also significant for the 2 varieties thus studied, namely "Caturra" and "Bourbon",

although the main effect of variety was not significant in this particular case.

AMÉRICO GROSZMANN (Serviço Nacional de Pesquisas Agronômicas, Rio de Janeiro). Ministerio da Agricultura

404 Growth Rate Studies in Corn.

Growth differences in ten day intervals between two inbreds and their reciprocal hybrids were studied in a split-plot experiment with three replicates, having ten dates of cutting. Each plot consisted of five single plant hills. Growth curves are difficult to analyze statistically. Several methods were tried to fit the basic purpose of the present study. R. A. Fisher in *Statistical Methods for Research Workers* describes the percentage relative growth rate method which seemed to give the best answer for the analysis of the differences. In the first growth period there were no significant differences between inbreds, F_2's and backcrosses. The F_1's exhibited a significantly lower growth rate. The inbreds grew most rapidly during this period.

FREDERICO PIMENTEL GOMES (Escola Superior de Agricultura "Luiz de Queiroz", Universidade de S. Paulo).

405 The Analysis of Factorial Experiments in Balanced Incomplete Blocks.

The author explains in detail how factorial experiments can be set up in balanced incomplete block designs, which permits the reduction of block size without confounding.

The analysis can be carried out with the adjusted means, by the same procedure used for factorials in complete block experiments, but afterwards the sum of squares thus obtained should be multiplied by a correction factor. This factor is $\lambda v/kr$ for the case of intra-block analysis, where v is the number of treatments and the other letters have the usual meaning.

When the recovery of inter-block information is carried out, the correction factor turns out to be

$$\frac{\lambda v + (r - \lambda)\hat{a}}{kr},$$

where,

$$\hat{a} = \frac{(r - 1)V_r}{rV_b - V_r}$$

V_r is the residual variance and V_b is the mean square for blocks (adjusted), and for Cochran and Cox's type I balanced incomplete block

design,

$$\hat{a} = \frac{[v(r-1) - k(g-1)]V_r}{k(b-g)V_b - (v-k)V_r}$$

for type II (g is the number of groups of replications present in the design used) and

$$a = \frac{v(r-1)V_r}{k(b-1)V_b - (v-k)V_r}$$

for type III.

An example of analysis of a $2 \times 2 \times 2$ factorial in balanced incomplete block, with $v = 8$ treatments and $b = 28$ blocks of $k = 2$ plots per block is given.

Papers presented at joint Biometric Society (ENAR) and I.M.S. sessions, Washington, March 7–9, 1957.

406 FORMAN S. ACTON. **The Mutual Difficulties of Statisticians and Digital Computers.**

The mutual difficulties of statisticians and digital computers arise from troubles in communication and a lack of generality in the formulation of statistical problems. The effort of programming and coding a problem for a digital computer is sufficient to require that a class of problems be included in any efficient routine. Statistical problems are still so heterogeneous as to resist this compacting.

Machines obey languages that are unsuited to the accurate programming of complicated indicial arithmetic. Most statistical problems fall into this category. The obvious cures are suggested, with some guarded optimism being expressed.

407 G. E. P. BOX (Statistical Techniques Group, Princeton University, Princeton, New Jersey). **Iterative Experimentation.**

Scientific research is usually an iterative process. The cycle: conjecture-design-experiment-analysis leads to a new cycle of conjecture-design-experiment-analysis and so on. It is helpful to keep this picture of the experimental method in mind when considering statistical problems. Although this cycle is repeated many times during an investigation, the experimental environment in which it is employed and the techniques appropriate for design and analysis tend to change as the investigation proceeds.

Broadly speaking, one or more of the following four phases can be detected in most investigations:

(a) a screening phase in which an attempt is made to isolate the important variables;

(b) a descriptive phase in which the effects of the variables and the positions of interesting regions of the space of the variables are empirically determined;

(c) a phase leading from (b) to (d);

(d) a theoretical phase in which an attempt is made to understand the actual mechanism of the process studied.

The roles which statistical methods should properly play in assisting the iterative process at these various phases of experimentation were briefly discussed.

408 J. E. FREUND AND W. O. ASH. **Some Estimation Problems in Generalized Harmonic Analysis.**

An important problem in the analysis of stationary ergodic Gaussian processes is the estimation of the power spectral density function, $\Phi(\omega) = 2/\pi \int_0^\infty \rho(\tau) \cos \omega\tau \, d\tau$, where $\rho(\tau)$ is the autocovariance function. The classical treatment of this problem is to estimate first $\rho(\tau)$ for various values of τ by systematic subsampling the process and then using numerical integration for an estimate $\hat{\Phi}(\omega)$. While the classical estimator can be shown to be biased, it has nevertheless proved to be adequate in many applications. It has the disadvantage, however, of requiring a sizeable number of numerical calculations in order to produce a single estimate of $\Phi(\omega)$.

An effort to get an estimator substantially as good as $\hat{\Phi}(\omega)$ but requiring considerably less numerical calculation has led to a new estimator for the spectral density,

$$\Phi^*(\omega) = \frac{1}{n} \sum_{i=1}^{n} x(t_i)x(t_i + k_i\Delta t)G(k_i),$$

where the k_i are independent random variables having equal probability distributions $P(k_i)$ and the $G(k_i)$ are weight functions. It is shown that the bias of $\Phi^*(\omega)$ can be made the same as the bias of $\hat{\Phi}(\omega)$ by suitable choice of $G(k_i)$ and $P(k_i)$ and that the difference in the variances of the estimators may be expressed as

$$\text{Var } \Phi^*(\omega) - \text{Var } \hat{\Phi}(\omega) = \frac{1}{n} \sum_{j=1}^{n} [\rho^2(0) + \rho^2(j\Delta t)] K^2(j) \left[\frac{1}{P(j)} - 1 \right],$$

where $K(j) = G(j)P(j)$. It was also discussed how to minimize the variance by a suitable choice of the probabilities and weights.

409 D. G. HORVITZ, G. T. FORADORI, J. MONROE, J. FLEISCHER AND A. L. FINKNER (North Carolina State College, Raleigh, N. C.) **An Investigation on the Smoking Habits of Individuals.**

An investigation was undertaken in cooperation with the American Tobacco Company to determine the most efficient method of measuring the amount of average daily smoking of individuals. This study was conducted in three phases.

Phase I

(a) Several semi-objective techniques for measuring average daily current smoking were tested. A counter attached to the case of a Zippo lighter, which is activated when the lid of the lighter is closed, was adjudged the most accurate among those tested.

(b) Approximately 92% of the variability among individuals in average weight of cigarettes smoked per day is explained by average number of cigarettes smoked per day. An additional significant proportion of the total variability is explained by using separate regressions for each type of cigarette, i.e. 70 mm, 85 mm and 85 mm filter.

Phase II

Several questionnaire measures were compared with the semi-objective technique selected as the standard. The questionnaire having the following classification:

(i) less than 10
(ii) 10–20
(iii) 21–40
(iv) over 40

is improved by rearranging the intervals so that the values 10, 20, 30 and 40 cigarettes smoked per day occur at some point within the interval rather than at the end of the interval.

For estimating average smoking per day for a specified group, there is some statitistical evidence favoring a questionnaire measure obtained by weighting the number smoked yesterday (a week day) by 5, number smoked Saturday by 1 and number smoked Sunday by 1. From a practical standpoint one question asking the number smoked, on-the-average may be equally as efficient. With the amount of data available, no difference in rates of misclassification could be detected among the questionnaire measures.

Phase III

One-half of the permanent factory and stemmery employees of the American Tobacco Company were to be administered the questionnaire used by the Census Bureau in the 1955 Current Population Survey, and one-half were to be given a new questionnaire developed by the Institute of Statistics. Departures from this 50:50 ratio were due to random absenteeism. The completion rate of the schedules was 98.1%. Although it could not be determined which was the most accurate, there appeared to be real differences in the classification of individuals by the three questionnaire measures considered.

410 CARL F. KOSSACK (Purdue University). **A New Approach To General Purpose Sampling.**

A sequential design procedure for introducing several variables into a survey design so as to have the final design meet the individual requirements on each variable is considered in the case where a stratified sampling plan is used in which one stratum is a census. The problem of developing a population list which meets the coverage requirements for each variable is first resolved by preparing successive listings of the sampling units ordered in turn by each variable and tagging the units needed to meet the respective variable coverage requirement. Tagged variables are always kept at the top of subsequent listings. At the end all tagged units are kept in the population. In designing the sample, the census stratum is built up successively so as to assure that the accuracy requirement for each variable is met, using estimated sampling rates for the other strata. For the final variable an optimum computational design is made using the built-up census stratum. If the sampling rates thus obtained differ significantly from the original estimates, a convergent process of design is recommended. The application of the procedure to a two-strata cluster sampling plan is discussed in detail.

411 H. L. LUCAS (North Carolina State College and Princeton University). **Some Uses of the I.B.M. 650 in Applied Statistics.**

Experiences of the author and his colleagues with the I.B.M. 650 seem to fall into three classes:

(i) analysis of data,
(ii) evaluation of unwieldy expressions including iterative solutions of complicated equations, and
(iii) empirical investigation of sampling distributions.

In analysis of data difficulties arise from memory limitations, diversity of types of data and desired final results. Some procedures used to partition large or complicated problems and to meet the varied demands were described. In class (ii) is an iterative scheme for finding the points of non-symmetrical rotatable designs and a method to arrive at the expectations of complicated quartic and higher forms. The latter was outlined. Empirical studies have been made or are in progress on genetic advance under selection with chromosome crossover, certain sequential procedures, a life testing problem, some facets of non-linear estimation, and Type I Type II errors in the unweighted means analysis of disproportionate data and the chi-square contingency test. Features of these familiar to the author were related.

412 DALE M. MESNER (Purdue University Center, Fort Wayne, Indiana). **The Structure of Incidence Matrices of Partially Balanced Incomplete Block Designs.**

In a *PBIB* design with v treatments and s associate classes, define $v \times v$ incidence matrices $\mathbf{A}_i = (a^i_{\mu\nu})$, $i = 1, \cdots, s$; $a^i_{\mu\nu} = 1$ if treatments μ and ν are i-th associates, $a^i_{\mu\nu} = 0$ otherwise. These matrices were introduced by Bose. Their properties and several applications, some of them new, are reviewed here. Some of the applications were presented by the writer at an earlier meeting (*Ann. Math. Stat. 27* [1956] 1185). The association schemes of designs of Latin square type with g constraints are defined by certain square arrays and have the property that fairly large sets of treatments are pairwise first associates. The matrices $\mathbf{A}_i$ are used in a proof that for fixed g and sufficiently large v this property follows from the values of n_i and p^i_{jk} and in turn implies the structure of the square array. This constitutes a uniqueness proof of these schemes.

413 M. A. SCHNEIDERMAN, R. J. TAYLOR AND S. B. FAND (National Institutes of Health, Public Health Service, Bethesda, Maryland). **Some problems in the Clinical Trials of Anti-Cancer Agents.**

The design of clinical studies to evaluate anti-cancer agents requires:

a. sharp definition of the disease studied
b. recognition of the special nature of advanced cancer
c. recognition of the problems of associated toxicity, and the problems of "cost" in toxicity as a negative measure of the effectiveness of a material. The concept of the therapeutic ratio is discarded in favor of a response-for-a-given cost. The

approach permits the evaluation of combination therapy, while the therapeutic index does not.

Problems of measurement are discussed and a presentation is made of the results of

a. measurement by different individuals
b. the handling of serially correlated data (which all tumor measurements are) from a maximum likelihood point of view, instead of the common least squares approach which assumes independence.

Sample charts of the "cost" approach, of examples of measurement problems, and a "Protocol Planning Guide" are given.

RAYMOND E. VICKERY. (Agr. Est. Div., U.S. Department
414 of Agriculture.) **Recent Experiences with Area Sampling for Agricultural Statistics.**

Probability area-sample surveys were conducted by the Agricultural Estimates Division in June of 1954, 1955 and 1956 as part of an extensive research program aimed at developing an improved crop and livestock estimating system. They were conducted under practical operating conditions to provide a better indication of how those procedures would behave when put to use on a large scale and also to train the nucleus of a staff for large scale operations. The program covered 700 segments of about 4 farms each in the 10 Southern States in 1954 and 1955. It was expanded to include 13 North Central States in 1956.

The approach in 1954 was to enumerate every farm for which the operator resided in the segment. There was heavy over-reporting of certain crops, especially those grown on a share basis. In 1955 the questionnaire was designed to eliminate this over-reporting. A list of 1,000 "large" farms was also added to the area sample that year. Although the over-reporting was corrected in 1955, sampling errors were not reduced appreciably.

In 1955, a separate sample of 100 segments was used to test the practicability of accounting for the use of all land falling wholly within the boundaries of sample segments with respect to both crop acreage and livestock inventories. This "closed-segment" approach appeared so satisfactory that it was used in the entire survey covering 1100 segments in 23 States in June 1956. Results clearly indicate that land use items can be estimated accurately by the closed segment method provided the sample is of sufficient size. Further studies are being made of the feasibility of gathering livestock data on closed segments.

On the average, one closed segment yields at least as much information as two segments using the farm headquarters approach.

415 R. LOWELL WINE. **A Sampling Study of Sources of Information for Farm Families in Virginia.**

A statistical study was conducted for the Virginia Agricultural Extension Service to determine which of its services, along with all other media of communication, have been most effective in reaching and helping the farmer with regard to two practices in the home. The "area sampling method" was used in the "open country" some within the West Central District of Virginia which includes sixteen counties. Two per cent of all farm families were interviewed in order to determine the source (sources) concerning the variety of his main field crop and type of fertilizer used on this crop and the homemaker furnishing the answers relative to her source (sources) concerning any major kitchen improvement and method of preserving foods for home consumption.

Some of the advantages and difficulties encountered in such a study are brought out and in some cases suggestions are made for improving the questionnaire.

THE BIOMETRIC SOCIETY

Australasian Region

Among the papers given at the 32nd meeting of the A.N.Z.A.A.S., Dunedin, 16–23 January, 1957 were: *Statistics Symposium*, G.S. Watson, "Chi-squared test for the goodness-of-fit of normal distributions"; P. Whittle, "Non-linear stochastic processes"; J. H. Darwin, "Some models of population growth"; W. M. Harper and J. A. Macdonald, "Distribution of the mean half-square successive difference in sampling from a normal population"; *Statistical Genetics Symposium*, A. H. Carter, "On estimating heritability"; B. I. Hayman, "A representation of gene action".

Brazilian Region

The Brazilian Region of the Biometric Society held its 3rd meeting at the Instituto Biologico, Sao Paulo, January 15, 1957. Scientific sessions were held at 10 a.m. and 4 p.m. Papers given included R. A. da Silva Leme, On checking the presence of a positive lower limit of a distribution of tensile strengths through an incomplete block design; F. Pimentel Gomes, Analysis of a group of thirty-eight fertilizer experiments with sugar-cane; P. Mello Freire and M. Picosse, Use of normalizing transformation in the analysis of anthropometric indices; C. G. Fraga, Jr. and R. Meirelles de Miranda, Covariance in a nonorthogonal experiment; A. Conagin, Estimation of the number of repetitions to be made in future experiments; A. Groszmann and Venita S. Nascimento, Variance components in the interpretation of a series of biological data; F. Pimentel Gomes, Elementary proof of Scheffé's test; and F. G. Brieger, Analysis of contrasts.

The annual business meeting was held at 3 p.m. Officers were reelected for the 1957 term: President—C. G. Fraga, Jr., Secretary—P. Mello Freire, Treasurer—A. Groszmann. Commission members elected in 1957 for three year term are A. M. Penha and A. Conagin; elected in 1956, F. G. Brieger (1956), L. Freitas Bueno (1956), J. M. Pompeu Memoria (1956, 1957) and A. A. Bitancourt (1956, 1957).

British Region

At a meeting on March 7, 1957, the following papers were presented: M. R. Sampford, A linearly balanced design for dairy cattle experiments; G. G. Meynell, The inherently low precision of infectivity assays; G. A. Barnard, Why fix totals in tests on 2×2 contingency tables?

ENAR

A joint meeting with the Institute of Mathematical Statistics was held in Washington, D. C. during March 7–9, 1957. Attendance at many sessions was upwards of 200. Among the papers presented were: A. Sample Survey Methodology: L. Wine, Sources of information for farm families in Virginia; C. Kossack, General purpose sampling; R. E. Vickery, Area sampling for agricultural statistics. B. Stochastic Processes: J. E. Freund and W. O. Ash, Generalized harmonic analysis; F. W. Diederick, Applications to aeronautic problems. C. Design of Experiments: W. H. Horton, Fractional factorials in industry; G. E. P. Box, Problems in evolutionary operation; C. Y. Kramer, Factorial treatments in group divisible incomplete block designs; D. M. Mesner, Structure of incidence matrices of PBIB designs. D. Electronic computers: H. O. Hartley, Reduction in programming by standardized routines; H. L. Lucas, Some uses of the IBM 650 in applied statistics; F. S. Acton, The mutual troubles of statisticians and digital computers. E. Survey of Smoking Habits: D. G. Horvitz, G. T. Foradori, J. Monroe, J. Fleischer and A. L. Finkner, Smoking habits of individuals. F. Statistics and Probability: J. Cornfield, Statistical inference; G. Noether, Nonparametric tests; E. Lukacs, Analytic characteristic functions; N. Severo, Tests of the means of certain distributions. G. Public Health and Medical Statistics: M. G. Sirken, Collecting data from samples of recently deceased persons; M. A. Schneiderman, R. J. Taylor and S. Fand, Clinical trials of anti-cancer agents.

Région Française

A l'Assemblée Générale, tenu le 13 février 1957, M. Michel Ollagnier a été élu au Conseil. M. André Vessereau a présenté une communication entitulée "Sur les conditions d'application du criterium χ^2 de Pearson".

Italian Region

A second course of Biometric Methodology was held at the Istituto Sierterapico Milanese from October 8 to 20, 1956. The program was similar to that of the first course held at Varenna the previous year. The following courses of lectures were given by L. L. Cavalli-Sforza, G. A. Maccacaro, R. Scossiroli and F. Sella: (1) Applied Statistical methods (11 lectures), (2) Design of experiments (5 lectures), (3) Bioassay (4 lectures). Each lecture was followed by a half hour discussion and one and a half hours practical work with the students working in pairs. Further theoretical courses were (4) Health statistics (A. Tizzano), (5) Theoretical foundations (F. Brambilla), (6) Clinical statistics (G. Barbensi).

Most of the students were graduates in medicine, but some attended whose main interests lay in veterinary medicine, general biology, pharmacy and other fields. Of 60 applications only 32 could be accepted. It is planned to hold a similar course during 1957.

Members

The following notifications of changes of address and of location of new members were received during February-April, 1957.

New Addresses

Dr. Peter Armitage, Cancer Chemotherapy, National Service Center, National Institutes of Health, Bethesda 14, Md., U.S.A.

Dr. Geoffrey Beall, Division of Manufacturing, Gillette Safety Razor Co., Boston 6, Mass., U.S.A.

Dr. F. E. Binet, Poultry Research Centre, Tarneit Road, Werribee, Victoria, Australia

Dr. Archie Blake, 2133 N. Circle Drive, Ann Arbor, Michigan, U.S.A.

Mr. James A. Bond, Dept. of Zoology, University of Chicago, Chicago 37, Illinois, U.S.A.

Mr. A. G. Constantine, C.S.I.R.O., Division of Math. Statistics, University of Adelaide, Adelaide, South Australia

Robert Joseph Deam, B.P. Ltd., Beaufort House, Gravel Lane, London, England

Mr. R. de Coene, I.N.E.A.C., Bambesa, Buta, Belgian Congo

Alison Grant Doig, c/o Professor M. G. Kendall, London School of Economics, Houghton St. & Aldwych, London W.C. 2, England

Mr. George E. Ferris, Apt. 46V, Building 8, 177 White Plains Road, Tarrytown, N. Y., U.S.A.

Mr. Willis W. Frankhouser, Merck, Sharp and Dohme, West Point, Pennsylvania, U.S.A.

Dr. J. M. R. Franckson, 21 Dieweg, Uccle, Belgium

Dr. P. W. Geier, c/o Australian Scientific Liaison Office, Africa House, Kingsway, London, W.C.2, Great Britain

Dr. Benson Ginsburg, Cobb Hall 215, University of Chicago, Chicago 37, Illinois, U.S.A.

Dr. Mordecai H. Gordon, Box 546, Perry Point, Maryland, U.S.A.

Dr. Theodore H. Greiner, M.D., Anderson Dept. of Psychiatry, Baylor University College of Medicine, Houston, Texas, U.S.A.

Dr. J. W. E. Harrisson, Library, P.C.P. and S., 43rd and Kingsessing Avenue, Philadelphia 4, Pennsylvania, U.S.A.

Brian Ivanhoe Hayman, Crop Research Division, Private Bag, Christchurch, New Zealand

Mr. J. A. Heady, Social Medicine Research Unit, Research Laboratories, London Hospital, Ashfield Street, London E. 1, England
Mr. L. Harmon Hook, Apt. 23, 5427 University Avenue, Chicago 15, Illinois, U.S.A.
Mr. Jay D. Leary, Jr., 70 Highland Street, Reading, Massachusetts, U.S.A.
Dr. P. H. Leslie, Bureau of Animal Population, Botanic Garden, High Street, Oxford, England
Mr. D. F. Matzinger, Department of Experimental Statistics, Box 5457, North Carolina State College, Raleigh, N. C., U.S.A.
Miss Jean Miller, 12 Mill Lane, Cambridge, England
Professor Per Nylinger, Skoghogskolan, Stockholm, Sweden
Mr. Floyd R. Olive, U.S.O.M., c/o American Embassy, La Paz, Bolivia
Dr. Bronson Price, 5813 Temple Hills Road, Washington, D.C., U.S.A.
Mrs. Lila Knudsen Randolph, 8004 Riverside Drive, Cabin John Park, Maryland, U.S.A.
Ingénieur Norbert Roussel, 15 Rue Combattants, Tirlemont, Belgium
Dr. V. Sahleanu, Schitu Meguresnuss, Bucharesti, Rumania
Professor Folmer D. Smith, Buoy, Tregde, Norway
Ph van Riessen, Cornelis Jostraat 115, Scheveningen, Netherlands
Mr. Bruno J. Vildosola, Sub-departmento de Bioestadistica, Casilla 3979, Santiago, Chile
Mr. William Weiss, Box 232, R.F.D. 4, Vienna, Virginia, U.S.A.
Professor Max A. Woodbury, College of Engineering, New York University, New York 53, N. Y., U.S.A.

New Members

At-Large

Professor Dr. Alexander Alexandrovich Lubischew, Uljianovsk, Krasnoarmeiskaya St. 2, K 4, U.S.S.R.
Mr. G. E. Hodnett, Regional Research Centre, The Imperial College of Tropical Agriculture, St. Augustine, Trinidad, B.W.I.
Professor Dr. Maximo Valentinuzzi, Calle Gascon 520, Buenos Aires, Suc. 13, Argentina, South America

Australasian

Dr. A. H. Carter, 88D Peachgrove Road, Hamilton, New Zealand
Mr. M. L. Dudzinski, C.S.I.R.O., Box 109 City, Canberra, A.C.T., Australia
Mr. C. R. Heathcote, Statistics Department, University of Melbourne, Carlton N.3, Victoria, Australia
Professor A. L. Rae, Massey College, Palmerston North, New Zealand

Brazilian

Dr. Alfredo C. Nascimento Filho, Caixa Postal 25, Rio de Janeiro, Brazil

Dr. Eneas Salati, Escola Superior de Agricultura, Piracicaba, Estado de Sao Paulo, Brazil

Belgian

Marcel J. W. Luttgens, Uangambi, B. P. 37, Belgian Congo

British

F. W. Bodmer, Dept. of Genetics, 44 Storey's Way, Cambridge, Great Britain

Dr. C. Daly, Glaxo Laboratories, North Lonsdale Rd., Ulverston, Lancashire, England

J. S. Gale, Dept. of Genetics, 44 Storey's Way, Cambridge, Great Britain

Dr. D. Lindley, Statistical Laboratory, St. Andrews Hill, Cambridge, England

Dr. C. C. Spicer, Central Public Health Laboratory, Collindale Avenue, London N.W. 9, England

ENAR

W. P. Cortelyou, Chairman, Department of Chemistry, Roosevelt University, Chicago 5, Illinois, U.S.A.

Charles F. Federspiel, Dept. of Biostatistics, University of North Carolina, Drawer 229, Chapel Hill, N. C., U.S.A.

Mr. James Grizzle, 11 B Davie Circle, Chapel Hill, N. C., U.S.A.

Professor John Gurland, Statistical Laboratory, Iowa State College, Ames, Iowa, U.S.A.

Dr. M. Hansen, Bureau of the Census, Washington, D.C., U.S.A.

Mr. Arthur G. Itkin, 925 Jersey Avenue, Elizabeth, New Jersey, U.S.A.

Dr. Benjamin Pasamanick, Research Division, Columbus Receiving Hospital, O.S.U. Health Centre, Columbus 10, Ohio, U.S.A.

Mrs. Mary E. Ready, Route No. 5, Frederick, Maryland, U.S.A.

Dr. Richard W. Roberts, 837 First Street, Rothschild, Wisconsin, U.S.A.

Dr. A. Sprott, 167 Glen Road, Toronto, Canada

Mrs. Hanna D. Sylwestrowica, 153 Park Avenue, Madison, N. J. U.S.A.

WNAR

Stanley R. Hill, Metabolic Lab., College of Osteopathic Physicians

and Surgeons, 1721 Griffin Avenue, Los Angeles 31, California, U.S.A.

French

M. René Chouchan, Ingénieur en chef du Service Statistique, Compagnie Française des Métaux, Givet, (Ardennes) France
Monsieur Maurice Fautrel, 8 Avenue Alphand, Paris 16, France
M. Guy Roberty, Institute d'Enseignement et de Recherches Tropicales, Route d'Aulnay, Bondy, (Seine), France

German

Prof. Dr. Ernst Assmann, Waldecker Hohe 130½, (13b) MIESBACH/ Oberhayern, Germany
Dr. Detlev Bruning, 19 Stendal, Tangermunderstr. 3b, Fernruf 684, Germany
Gunther Caroli, Freiburg Br., Bertholdstr. 17, Germany
Dr. Heinz Fink, Ludwigshafen am Rhein, Pfalxfrafenstr. 46, W. Germany
Dr. Gerhard Specht, Kleinmachnow b, Potsdam, Philipp-Muller-Allee 42, Germany
Dr. Friedrich Wasserfall, *Kiel*, Kronshagener Weg 32, Germany
Dr. Franz Weiling, Bonn, Rhein, Lengsdorf, Kapellenstrasse 65, West Germany

Italian

Dr. E. Robotti, Villaggio Sanatoriale de Soudale, (prov. Soudrio) Italy
Gian Tommaso Scarascia, Viale Mazzini 13, Rome, Italy

Netherlands

Dr. J. A. H. Gooszen, Dolderseweg 158A, Den Dolder, Netherlands
Dr. Franz Adolf Nelemans, Cornelis Houtmanstraat 18, Utrecht, Netherlands
Mr. H. A. Tas, Banstraat 59b, Amsterdam 21, Holland
Jan T. N. Venekamp, Eelderwelde 13, Groningen, Post Eelde, Netherlands
J. C. A. Zaat, Wageningen, Dr. Wiemayer Str. 6, Holland

Swiss

Dr. Christian Auer, Lurlibadstr. 115, Chur, Switzerland
Dr. Edgar Grasemann, Institut fur Haustierernahrung des Eidg, Technischen Hochschule, Universitatstr. 12, Zurich, Switzerland
Prof. Dr. Hans Lortscher, Animal Breeding Institute, Swiss Federal Institute of Technology, Universitatstr. 2, Zurich, Switzerland
Martin Menzi, Morgensalstrasse 21, Zurich 38, Switzerland

THE BIOMETRIC SOCIETY

Secretary's Imprest Account

Statement of Income and Expenditure during the period ended 31st Dec., 1956.

Income	£	s.	d.	£	s.	d.
Treasurer—$600 Imprest				213	4	11
Membership Subscriptions (2)				3	3	8
				216	8	7
less Expenditure						
Office Furniture	43	6	7			
Stationery	37	10	10			
Secretarial assistance	31	5	—			
Postages	4	3	10			
Sundries	3	4	3	119	10	6
Balance in Hand—31.12.56.				£96	18	1

I certify the above to be a true record of my transactions on behalf of the Biometric Society.

12th March, 1957 — M. J. R. Healy (Signed)
Secretary

I have examined the account book and vouchers produced by the Secretary and certify that the above statement is in accordance therewith.

12th March, 1957 — E. Church (Signed)
(E. Church) A.A.C.C.A.

Treasurer's Report, 1956

Balance Sheet

Assets		
Cash—$10,031.18 less	$7,441.22	$2,589.96
(Bank balance	$2,554.96	
plus Petty Cash	$ 35.00)	
Liabilities		
Subscriptions, 1957	$ 29.75	
Dues, 1957	20.25	$ 50.00
Surplus, 1/1/56	$1,956.89	
Gain for Period	583.07	2,539.96
		$2,589.96

Income and Expenditure Statement

INCOME		
Subscriptions—1955	$ 473.50	
1956	4,178.50	$4,652.00
Dues —1955	$ 137.25	
1956	1,973.75	2,111.00
Sustaining Memberships, 1956		455.00
Back dues and subscriptions		105.75
Regional allotments	$ 78.50	
BIOMETRICS allotments from Sustaining Members	175.00	
Back issues	247.50	
Reprints and directories	3.40	
Journal of A.S.A. for Biometric Society Members	115.00	
Striplist	10.00	
Sale of furniture and equipment	219.00	
Consular fee	3.00	
Bank charges	.25	
Overpayments	72.17	
	$ 923.82	
Less credits and allotments used	214.28	709.54
Total Income		$8,033.29
EXPENDITURE		
Secretary's office	$ 609.00	
Subscription to ASSOCIATIONS	5.00	
Subscriptions to J.A.S.A. for Society Members	120.00	
Shipping and trucking charges	110.33	
Consular fee	3.00	
Bank charges	.58	
Miscellaneous	.60	
BIOMETRICS	4,897.11	
Special services	78.50	
Postage	95.85	
Stationery and supplies	72.83	
Telephone	85.83	
Printing	70.93	
Salaries	1,300.66	
Total Expenditures		$7,450.22
Excess of Income over Expenditures		583.07
		$8,033.29

Audited: Charles A. Smith

Date: January 14th, 1957.

Operations Statement of Biometrics Volume 12 (1956)

INCOME (1 February 1957)			
Subscribers			
Biometric Society			
557	$ 4.00	$ 2,228.00	
734	2.75	2,018.50	
7 Sustaining	25.00	175.00	$ 4,421.50
493 ASA	4.00		1,972.00
759 Direct	7.00		5,313.00
Sale of back issues			2,495.97
Sale of reprints			692.94
Exchange			.87
			$14,896.28
EXPENDITURES (1 February 1957)			
Cost of Journals			
Printing			
Issue No. 1	$ 1,526.23		
Issue No. 2	2,190.59		
Issue No. 3	1,946.67		
Issue No. 4	3,257.60	$ 8,921.09	
Mailing and Express Charges			
Issue No. 1	134.58		
Issue No. 2	163.68		
Issue No. 3	143.75		
Issue No. 4	221.18	663.19	9,584.28
Cost of Reprints			
Printing			
Issue No. 1	163.44		
Issue No. 2	324.77		
Issue No. 3	169.02		
Issue No. 4	343.73	1,000.96	
Mailing Charges			
Issue No. 1	11.63		
Issue No. 2	27.26		
Issue No. 3	15.83		
Issue No. 4	26.59	81.31	1,082.27

OPERATING EXPENSES		
Stamps	$ 330.00	
Stationery	133.43	
Duplication Work	52.02	
Telephone	3.75	
Insurance	136.03	
Customs	5.63	
Bank Charges	15.82	
Exchange—Transfer of cheques	22.25	
Joseph Ruzicka, Book Binding	19.35	
Express Charges	28.52	$ 746.80
		$11,413.35
Income		14,896.28
Expenditure		11,413.35
Surplus		$ 3,482.93

Balance Sheet Biometrics Volume 12

ASSETS (1 February 1957)			
Accounts Receivable		$ 590.16	
Bank Balance			
U.S.	$ 7,491.05		
Canadian	1,085.33	8,576.38	
U.S. Treasury Bond		5,000.00	
LIABILITIES (1 February 1957)			
Subscriptions to Volume 13			$ 1,491.00
Balance from Previous Volumes			8,602.45
Surplus from Volume twelve—			
Including Accounts Receivable			4,073.09
		$14,166.54	$14,166.54

Audited:
A. W. Quealy

NOTE:
Not included in Assets is Stock of back issues from Volume 1–12 and Reprints.

NEWS AND ANNOUNCEMENTS

Members are invited to transmit to their National or Regional Secretary (if members at large to the General Secretary) news of appointments, distinctions or retirements and announcements of professional interest.

D. R. Cox of the University of North Carolina has been appointed to a Readership in Statistics, University of London at Birbeck College.

Jack Moshman has left Bell Telephone Laboratories, Inc., where he was consulting statistician to assume the post of Director of the Division of Mathematical and Statistical Services of the Council for Economic and Industry Research, Inc.

D. E. W. Schumann, Head, Department of Statistics, University of Stellenbosch, South Africa and *R. A. Bradley*, Professor of Statistics, Virginia Polytechnic Institute, Blacksburg, Virginia, U.S.A. were co-authors of a paper entitled "The comparison of the sensitivities of similar experiments", which won the J. Shelton Horsley Research Award of the Virginia Academy of Science at the May meeting at Old Point Comfort, Virginia. The J. Shelton Horsley Research Award is awarded annually to the best paper submitted in the competition in Virginia.

M. C. K. Tweedie has accepted a position as Temporary Lecturer in Mathematical Statistics at the University of Manchester, England. He has been on the Faculty of the Virginia Polytechnic Institute, Blacksburg, Virginia, for the past four years.

Notices

Department of Statistics, University of Chicago

The department of statistics at the University of Chicago, known since its organization in 1949 as the Committee on Statistics, is now called the Department of Statistics. The name was changed from Committee to Department in order to avoid confusion about the nature and status of the organization.

Leonard J. Savage, who has been Acting Chairman of the Department this year, has accepted a regular appointment as Chairman beginning March 1, 1957. He succeeds *W. Allen Wallis*, who now is Dean of the School of Business though he continues as a member of this Department. Other members of the Statistics faculty are *K. A. Brownlee, Kai Lai Chung, Sudhish G. Ghurye, Leo A. Goodman, William Kruskal, John W. Pratt, Harry V. Roberts*, and *David L. Wallace*.

Research Center, General Foods Corporation

Central Laboratories, General Foods Corporation, 1125 Hudson Street, Hoboken, New Jersey, is changing its name and address to Research Center, General Foods Corporation, 555 South Broadway, Tarrytown, New York, effective July 1, 1957.

International Congress of Mathematicians 1958

At the invitation of the City and University of Edinburgh and the Royal Society of London, the International Congress of Mathematicians will meet in Edinburgh from August 14 to August 21, 1958. His Royal Highness the Duke of Edinburgh has graciously consented to extend his patronage to the Congress.

The Executive Committee is inviting a number of mathematicians to deliver one-hour and half-hour addresses. There will also be daily sessions devoted to fifteen-minute communications.

There will be eight sections, namely: 1. Logic and Foundations. 2. Algebra and the Theory of Numbers. 3. Analysis. 4. Topology. 5. Geometry. 6. Probability and Statistics. 7. Applied Mathematics, Mathematical Physics and Numerical Analysis. 8. History and Education. A program of entertainments and excursions is being planned.

Those who wish to receive further information about the Congress may write to Frank Smithies, Secretary of the International Congress of Mathematicians, Mathematical Institute, 16 Chambers Street, Edinburgh, 1, Scotland.

XVth International Congress of Zoology-London, 1958

The XVth International Congress of Zoology will take place in London from 16th-23rd July, 1958, under the Presidency of Sir Gavin de Beer, F.R.S., Director of the British Museum (Natural History) assisted by an Advisory Committee comprising all the leading British zoologists.

A number of special topics and sessions are being arranged which, while covering a wide field, will centre around Evolution. Congress will be organized into 12 Sections, and will be preceded by a Colloquium on the Rules of Zoological Nomenclature.

At the Inaugural Meeting in the Royal Albert Hall, Dr. Julian Huxley will give a special Darwin-Wallace Centenary address. At the concluding meeting, Professor J. Millot will speak on Coelacanths.

Further information may be obtained from the Registrar of the XVth International Congress of Zoology, c/o British Museum (Natural History), London, S.W. 7, England.